

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: October 21, 2001, 10:54:58 ; Search time 7124.02 Seconds  
(without alignments)  
11995.940 Million cell updates/sec

Title: US-09-515-806-1  
Perfect score: 5525  
Sequence: 1 tgcgccacggtccgcacc.....aatgtttcatataacctgca 5525

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 1344157 seqs, 7733874588 residues  
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba1.\*  
2: gb\_ba2.\*  
3: gb\_ba3.\*  
4: gb\_in1.\*  
5: gb\_in2.\*  
6: gb\_in3.\*  
7: gb\_om.\*  
8: gb\_ov.\*  
9: gb\_pat1.\*  
10: gb\_pat2.\*  
11: gb\_ph.\*  
12: gb\_pl1.\*  
13: gb\_pl2.\*  
14: gb\_pl3.\*  
15: gb\_pl4.\*  
16: em\_ba1.\*  
17: em\_ba2.\*  
18: em\_fun.\*  
19: em\_htgo\_hum.\*  
20: em\_htgo\_inv.\*  
21: em\_htgo\_rod.\*  
22: em\_htg\_hum1.\*  
23: em\_htg\_hum2.\*  
24: em\_htg\_hum3.\*  
25: em\_htg\_hum4.\*  
26: em\_htg\_hum5.\*  
27: em\_htg\_hum6.\*  
28: em\_htg\_hum7.\*  
29: em\_htg\_hum8.\*  
30: em\_htg\_inv1.\*  
31: em\_htg\_inv2.\*  
32: em\_htg\_other.\*  
33: em\_htg\_rod.\*  
34: em\_hum1.\*  
35: em\_hum2.\*  
36: em\_hum3.\*  
37: em\_hum4.\*  
38: em\_hum5.\*  
39: em\_hum6.\*  
40: em\_hum7.\*  
41: em\_in.\*  
42: em\_om.\*  
43: em\_or.\*

44: em\_ov.\*  
45: em\_pat.\*  
46: em\_ph.\*  
47: em\_pl.\*  
48: em\_ro.\*  
49: em\_sts.\*  
50: em\_sy.\*  
51: em\_un.\*  
52: em\_vi.\*  
53: gb\_sts1.\*  
54: gb\_sts2.\*  
55: gb\_sts3.\*  
56: gb\_sy.\*  
57: gb\_un.\*  
58: gb\_vil.\*  
59: gb\_vil2.\*  
60: gb\_htg1.\*  
61: gb\_htg2.\*  
62: gb\_htg3.\*  
63: gb\_htg4.\*  
64: gb\_htg5.\*  
65: gb\_htg6.\*  
66: gb\_htg7.\*  
67: gb\_htg8.\*  
68: gb\_htg9.\*  
69: gb\_htg10.\*  
70: gb\_htg11.\*  
71: gb\_htg12.\*  
72: gb\_htg13.\*  
73: gb\_htg14.\*  
74: gb\_htg15.\*  
75: gb\_htg16.\*  
76: gb\_htg17.\*  
77: gb\_htg18.\*  
78: gb\_htg19.\*  
79: gb\_htg20.\*  
80: gb\_htg21.\*  
81: gb\_htg22.\*  
82: gb\_htg23.\*  
83: gb\_htg24.\*  
84: gb\_htg25.\*  
85: gb\_pr1.\*  
86: gb\_pr2.\*  
87: gb\_pr3.\*  
88: gb\_pr4.\*  
89: gb\_pr5.\*  
90: gb\_pr6.\*  
91: gb\_pr7.\*  
92: gb\_pr8.\*  
93: gb\_pr9.\*  
94: gb\_rol.\*  
95: gb\_rol2.\*  
96: gb\_in4.\*  
97: gb\_pr10.\*  
98: em\_ba3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5156.4	93.3	5163	9	AX056426
2	4929.8	89.2	4994	85	AB037759
3	3909.8	70.8	5212	94	MM0243533
4	3896.4	70.5	5230	94	AF193343
5	3698.8	66.9	5322	94	AF193344
6	3418.4	61.9	5141	94	AF193342
7	3332	60.3	3351	93	HSN802494
8	2949.6	53.4	2986	89	AK027011
					AX056426 Sequence
					AB037759 Homo sapi
					AJ243533 Mus muscu
					AF193343 Mus muscu
					AF193344 Mus muscu
					AF193342 Mus muscu
					AL157497 Homo sapi
					AK027011 Homo sapi





Db 3306 ATGAGCACACGAGCTGCCCTATTTCATGACACAGCGGGATGCTGGTATGCTTCCTT 3365  
Qy 3409 ttgacctgagatcccttttgcaagatatgtgcaagaataataattgaattttaaac 3468  
Db 3366 TTGACCTGGCGATCCCTTTTGCAAGATATGTGCGCAGAAATATATATATGATTTAAAC 3425  
Qy 3469 gatactgcatagaacagtgtgttgcagccgcgaagttagatcgatttcacccaagaac 3528  
Db 3426 GATACTGCATAGAACGTGTCTTTCAGGCGCGCAAGTTAGATCGATTTCATCCCAAAGAAC 3485  
Qy 3529 ttctggaagtgcattgatattgtcaacttctaccaccaacagcttctgcccactgtcg 3588  
Db 3486 TTCTGGAGTGTGCATTGTGATTAATTTGCTCACTTCTACCAACAGCTTCTGCCACTGTCTG 3545  
Qy 3589 aaattatctacactatctatgaataatcatcaagatttccagcaactcaggaagaatt 3648  
Db 3546 AAATATCTACACTATCTATGAATCATCAAGAGTTTCCAGCACTTTCAGGAAGAAAT 3605  
Qy 3649 acagtattattgaaccataaccatgtatttgaagcaatactcttacactgtgggatcc 3708  
Db 3606 ACAGTATTATTGAAACCATACCATGTTATTGAAAGCAATACTCTTACACTGTGGGATCC 3665  
Qy 3709 cagaagaataactcaagtcgaatctacattattctgtatgatctgtgcagagaaactga 3768  
Db 3666 CAGAAGATAAATCAGTCAAGTCTACATTATCTGTATGATGCTGTGACAGAGAACTGA 3725  
Qy 3769 cgaggagagaagtgaagctaaattttgtaattctgtcttcttctcttaattgtctgtgc 3828  
Db 3726 CGAGGAGAGAGTGAAGCTTAATTTTGTATCTGTCTGTCTTCTTCTATAGTCTGTCTG 3785  
Qy 3829 gactctacaagttattgaacagaaggaatttgaagatcttattgccacaataaatt 3888  
Db 3786 GACTCTACAAGTTATTGACAGAGGGAGATTGCAAGATCTTATGCCAACATAAAT 3845  
Qy 3889 cattaataacagaaacaggattgtgcacagctgtgtgaagtatggtcttaagaacactag 3948  
Db 3846 CATTAATAAACAGAAACAGGATTTGCAAGTTGTGAGTATGGCTTAAAGACCTAG 3905  
Qy 3949 aggaaggtgtgactgtgaagaactcggcatcaagttacaggtcttgatcaattgg 4008  
Db 3906 AGGAGGTGTGTGACACTGTGAAGAACTCGGCATCAAGTTACAGTCTTGATCAATTTGG 3965  
Qy 4009 gcttggtttacaaggtgcagcagcaaatggaatcatcttccagtttggctttcatca 4068  
Db 3966 GCTTGTGTTTACAAGTGCAGCAGCACAAATGGAATCATCTTCCAGTTTGTGGCTTTCATCA 4025  
Qy 4069 aacgaagcaaaaggtgtaccctgaaatccctgcagctggagcagatatgaactctga 4128  
Db 4026 AACGAAGGCAAAAGGCTGTACCTGAATCTCTCGAGCTGGAGGAGATATGACCTGCTGA 4085  
Qy 4129 ttcccgagtttagagggccacaagctctggggccagttcccactgcattgggggtcagca 4188  
Db 4086 TTCCCCAGTTTAGAGGGCCACAGCTCTGGGGCCAGTTCCCACTGCCATTTGGGGTCAGCA 4145  
Qy 4189 tagctatagaagaatctctgctgtctcctcaacatggaaggaatctctgttacaataagct 4248  
Db 4146 TAGCTATAGACAAGATATCTGCTGTCTCTCAACATGAGGAATCTGTTACAAATAAGCT 4205  
Qy 4249 cttgtgacctctgtgttgaagttgtggtcagatgtctatgtccaggccatcaacctaa 4308  
Db 4206 CTTGTGACCTCTCTGTTGTAAAGTGTGGTCCAGATGTCTATGTCCAGGGCCATCAACCTAA 4265  
Qy 4309 ccsgaaactctgacagcaggcatcacagcagaaatcatgtacgactggtcacagctccc 4368  
Db 4266 CCCAGAACTCTGGACAGCAGGCATCACAGCAAAATCATGTACGACTGTGTACAGTCCC 4325  
Qy 4369 aagagggaattacaagactactgacagacatcatgaaatcaactatgtgcccctgtctcgg 4428  
Db 4326 AAGAGGAATTACAAAGAGTACTGCAGACATCATGAAATCACTATGTGGCCCTTGTCTCGG 4385  
Qy 4429 ataaagagaagccatgtcaaggttaagttcttctcagaaggaaggaagcagacagagaagc 4488  
Db 4386 ATAAGAGGAAGCCATGTCAAGTTAAGTCTTTCGAGGAAGGAAGGCAGACAGAGC 4445

Qy 4489 gtgtgtgagactgaactgttgaccatgtactgcagaaactgaggaactaaagtcaactg 4548  
Db 4446 GTGTGCTGGACACTGAACCTTCTGGACCATGTACTGCGAAACTGAGGACTAAAGTCACTG 4505  
Qy 4549 atgaagaagaatgagcagagaagcttccgataatttgcagtgcaaaatctaaaggggtcat 4608  
Db 4506 ATGAAGAAGAAATGGCAGAGAAGCTTCCGATAATCTTTCAGTGCAGAAATCTGAAGGGGTAT 4565  
Qy 4609 ttcttaagtctcaggtttgttgaatcccatgagcaaacagtggttcccatgtgagtg 4668  
Db 4566 TTTCTAATGCTTTCAGTGTGTTTGAATCCATGAGCAACAGTGTGTTCCCATTTGTGAGTG 4625  
Qy 4669 tctagccccggagaagctgtcagccagcactagggcgctatgaaactcagtagtaaaa 4728  
Db 4626 TGCTAGCCCCGGAGAAGCTGTGAGCCAGCACTAGGAGGCGCTATGAAACTCAGGTACAAA 4685  
Qy 4729 ctgacttcaagactcccttgcacacttacaatcagaagaagcagtgaaattgaaattctgg 4788  
Db 4686 CTGCACTTTCAGACCTCCCTTCCCACTTACATCAGAAAGCAGTGAATTTGAAATTTCTGG 4745  
Qy 4789 ctgtggtactaccacaaagaacaaatattcacagtttttatcattagagtggtggtctgatg 4848  
Db 4746 CTGTGGATCTTACCACAAAGAAATATTTACAGTCTTTCATCATTAGAGTGGATGCTCATG 4805  
Qy 4849 aacagcatttaacacaaactgtgaagcagctgtgcacgcctgcacgcctccaaagcaaaatacc 4908  
Db 4806 AACAGGCATTTAACACAACTGTGAAGCAGCTGCTGCACGCTGCCAAAGCAAAAGATACC 4865  
Qy 4909 teaatattagctgtgatgaatttataacatcaaaagtagaaaaaagggtctctgtgat 4968  
Db 4866 TCAAAATTAGTCTGTGATGAAATTTATAACATCAAGTAGAAAAAAGGTGCTGTGCTAT 4925  
Qy 4969 ttctacagctatagatgactactacagaactcttatttaacctaaagaactgtcg 5028  
Db 4926 TTCTGTACAGCTATAGACATGACTACTACAGACTTATTTTACCCCTAAAGACTGTGCG 4985  
Qy 5029 ttaacctcattcaaacagacagaggtcttactggaataatggaatgtgtgtacattcatc 5088  
Db 4986 TTAACCTCATTTCAACACAGACAGAGCTTATCTGGAATTAATGGAATGTGTACATTATC 5045  
Qy 5089 ataatttaaaatttaattcctaagaagagctggtgagctgcacaccttttaattccca 5148  
Db 5046 ATAAATTTAAATTTAAATTTCTAAGAGAGGCTGGGTGCGAGTGTGCTACACCTTTATCCCA 5105  
Qy 5149 gcactttgggaagccaaagcaggaagactgtctgaaccagagagtttgagaccagcct 5206  
Db 5106 GCACCTTTGGGAAGCCCAAGCAGGAGACTGCTTGAACCAAGAGTTTGAGACCAAGCCT 5163

## RESULT 2

AB037759  
LOCUS AB037759 4994 bp mRNA PRI 14-MAR-2000  
DEFINITION Homo sapiens mRNA for KIAA1338 protein, partial cds.  
ACCESSION AB037759  
VERSION AB037759.1 GI:7243056  
KEYWORDS  
SOURCE Homo sapiens brain cDNA to mRNA, clone\_lib:pBluescriptII SK plus clone:fh16948.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (sites)  
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.  
Prediction of the coding sequences of unidentified human genes.  
XVI. The complete sequences of 150 new cDNA clones from brain which  
code for large proteins in vitro

DNA Res. 7 (1), 65-73 (2000)

20181126

## MEDLINE

2 (bases 1 to 4994)

## AUTHORS

Ohara,O., Nagase,T. and Kikuno,R.

## TITLE

Direct Submission

## JOURNAL

Submitted (31-JAN-2000) to the DDBJ/EMBL/GenBank databases. Osamu













QY 1857 ggagcttttgagctgtcatcaaggtgagaaagttgagcaggtgtgctgtcagcagtg 1916  
Db 1884 GGAGCTTTGGAGGTGTCATCAAGGTGCAAAACAAGCTTGATGGCTGCTGCTATGCTGTG 1943  
QY 1917 aagcgatccccatcaaccggccagccggcagttccgagggatcaagggcgaagtgcaca 1976  
Db 1944 AAGGCGATCCCTATCAACCCCTGCAAGCAGACACTTCGCGCGGATCAAGGGCGAGGTGACA 2003  
QY 1977 ctgctgcaaggtgcaccatgagaacattgtcgcttactacaacgcctgtgagtcgagcgg 2036  
Db 2004 CTGCTATCGCCCTGACCATGAGAACATTTGCGCTACTTACAACGCCCTGATCGAGCGG 2063  
QY 2037 cagcagcgccggcgagccgggagccggcccccggactccgggccccttggcccaagggat 2096  
Db 2064 CATGAGGCCCGAGGGTGCAGGACACCGCCCGGACTGCACACCCCGAGGCCAGGAC 2123  
QY 2097 gaccgagctgcacggcgagccggcgagcagacacagcggcctggacagcgtlagagcc 2156  
Db 2124 AGCCGAGCCACTCGGGGAAACATCGGGGACACTGAAGAACTGGGCAGCGTGGAGGCT 2183  
QY 2157 gccgcgcgcacccactcctcagcagctcggtggagtgagcacttcggcgagcgctcg 2216  
Db 2184 GCAGCAGCCGCCACCATCTCAGCAGCTCGGTGGAGTGGAGCACATCTGCAGAGCGTTCT 2243  
QY 2217 gccagtgcccgtttcccgccacccggcccgccggggtccagcgatgacgagcagcag 2276  
Db 2244 ACCAGCACCCGCTTCCAGTCACGGGCCAGGATTCACG---AGCGATGAGGAGGAG 2300  
QY 2277 gacgagcagctggcgttctccagctcctcctcgtcgttccagcttctgaaagtgc 2336  
Db 2301 GACGAGCCGATGGGCTCTTCCCGAGTCTTCTACTGCTTCAGATCTTGACAGTGAC 2360  
QY 2337 attatctttgacaatgaagatgagaacagtaaaagtgcagaatcaggtgaagattgcaat 2396  
Db 2361 ATCATCTTTGACATGAAGATGAACAGTAAAGTCAAAATCAGGATGAAGAGTCAAT 2420  
QY 2397 gaaaagaatggctgcatgaaagtgcacatcagtcagcagcagcagcagcagcagcagc 2456  
Db 2421 CAAGAGGAGCGTAGGATGAGCTTTCAGTGAGCAGCTGAGGCTGTGCATACCTT 2480  
QY 2457 tacatccagatggagctgtgaaagagcactttagagacacacattgacacagggactg 2516  
Db 2481 TATATCCAGATGGAGTACTGCGAGAGAGACAGCTGCGGGACACATTTGACCGGGCTG 2540  
QY 2517 tatcagacacogtcagactctggaggcttttttccagagatctcggatggattgcttat 2576  
Db 2541 TTCCGAGACACCGCCGCTCTGGAGGCTTTTCCGAGAGATCTGGATGGATTAGCTTAT 2600  
QY 2577 atccatgagaagaatgattcacccggatgtgaagcctgtcaacatttttttgattct 2636  
Db 2601 ATCCATGAGAAAGGAATGATTCGCGACCTTGAAGCCCTGTCAATATTTTTTGGATTCT 2660  
QY 2637 gatgaccatgtaaaataggtgatttttgggttggcagacagcactctagccttttctgct 2696  
Db 2661 GATGACCATGTGAATAGGTGACTTTGGCCCTGGCAACAGACCATCTGGCCCTTCACTGCT 2720  
QY 2697 gacagcaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2753  
Db 2721 GAAGGTAAACAGGAGCAGCAGGAGTGACGGAGTGATTAAGTCAGACCCCTCAGGCCAT 2780  
QY 2754 ttaactggagtggtggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2813  
Db 2781 TTGACTGGCATGGTGGTACTGCTGTATGTATGTAAGCCCTGAGTCCAGGAGACCAAG 2840  
QY 2814 tctgcatacaaacagaaagtgtatcttccagcctgggaattatcttctttgagatgctc 2873  
Db 2841 TCTGCTACAAACAGAAAGTGGATCTTTCAGCCCTGGGAATTTATCTTCTTGGATGCT 2900  
QY 2874 tatccccatggctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2933  
Db 2901 TATCACCCGATGGTACTGCCCTCAAGAAAGATTTTCGTTCTCAACCAACTTCGAGATCCC 2960  
QY 2934 acttcgcctaaagtctccagaagactttgacgatggagcagcagcagcagcagcagcagc 2993

Db 2961 ACATGCCCAAGTTTCCAGACGACTTCGATGATGAGACATACAAAGCAGANAATCTGTC 3020  
QY 2994 atctctggtctgtgaaccacgataccagaaacgcccacagccacagaaactgctcaag 3053  
Db 3021 ATCTCTGCTGTTGAACCATGATCCAGCTAAACGCCACCGCCATGGAACCTGCTCAAG 3080  
QY 3054 agtgagctgctccccccccccccagatggagagtgagagtgatgaagtgcagagctgcacac 3113  
Db 3081 AGTGAGCTGCTCCCGCCCGCAGATGGAGAGTCCGAGCTGCATGAAGTGTCTGCCAC 3140  
QY 3114 acgctgacacacgctggatggagggcctaccgcacacatgagccagatctcttcgag 3173  
Db 3141 AGCTGCCCCAACATAGATGGGAGGCTTACCGCACCATGATGAGTCAGATCTTCTGCCAG 3200  
QY 3174 cgcactccccctgcctcagattacacactatgacagcagacatactgaagggcaacttctca 3233  
Db 3201 CACATCTCCCGCTATAGACTACACTACGACAGTGACATCTGAAGGGCAACTTCTTG 3260  
QY 3234 atccgtacagcaagatgcagcagcagctgtgtgaaccatcatccgcatctttaaaga 3293  
Db 3261 ATTCGTACAGCCCAAGATCCAGCAGCTTGATGTGAACCATCTGTCGGCTCTTTAAAGG 3320  
QY 3294 catggagctgttcagttgttactccactactgtctccccgcgaacacagacaaatatag 3353  
Db 3321 CATGGCTGCTCCAGCTGTGACCCCTGCTGCTTCCCGAAACAGGCAATATATGAG 3380  
QY 3354 cacaagagctgcccctatctatggaccacagcgagatgctggtgatgcttcttctgac 3413  
Db 3381 CACAAGAAAGGGCTTTGTTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 3440  
QY 3414 ctgcgactccctttgcaagatgtggcaagaaataatattgaatttaaagacatgac 3473  
Db 3441 CTGGGGTCCCTTTGCAAGATGATGGAAGAAATTAACATGATGATGATGATGATGATGAT 3500  
QY 3474 tgatagaaactgtgttcagggccgcaagtttagatcgatttccatccccaaagactctg 3533  
Db 3501 TGATAGAGGGGTGTTGACAGCTCGGAAGTTAGACCGATTTTCATCCCAAGAACTCCCTG 3560  
QY 3534 ggtgtgacttgcattgattgacttaccacacagcttttccgactgctgaaatt 3593  
Db 3561 GAATGTGATTTGATATGCTGCTTACACCAAGAGCTCTCTGCCACTGCGGAAACC 3620  
QY 3594 atcacactatcatgaaatcccaagagtttccagcacttcaggaagaaataacagt 3653  
Db 3621 ATCTACACCATCTATGAATATCCAGGAGTTTCTGCACTTCAGGAAAGGAATTACAG 3680  
QY 3654 attatttgaccatacctatgttattgaaagcaatactcttaccactgtgggattccagaa 3713  
Db 3681 ATTTACTTGAACCATACGATGCTTCTGAAGGCAATACTCTACACTGTGGGATCCGGAG 3740  
QY 3714 gataaactcagtcagctacatttctgtatgctgtgacagagcagcagag 3773  
Db 3741 GACAAGCTCAGCCAAAGTCTACGTCATTTCTGTATGATGCTGTGACAGAGAGCTGACTAG 3800  
QY 3774 agagaagtgaagcctaatttgaactcttcttcttcttcttcttcttcttcttcttct 3833  
Db 3801 AGAGAAGTGAAGCTAATTTCTGTAACTCTGCTTGTCTTCTTCTTCTTCTTCTTCTTCT 3860  
QY 3834 tacaagtttttgaacagagagagatgttgcagagcttcttgcacacaaataaattta 3893  
Db 3861 TACAAGTTTCTGAGCAGAGAGGGGATTTGCAAGACTTAAACGCCCAACTCACTACTA 3920  
QY 3894 ataaacagaaacaggtattgcacagctgtgagcttgaagcttgaagcacttgaagag 3953  
Db 3921 ATAAACAGAAACAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3980  
QY 3954 gttgtgactgttgaagaaactggcagctcaagttacagctcttgcataatttgggcttg 4013  
Db 3981 GTCTGCGGCTGCTGAGAAACTTGGTGAAGTTTACAGGTTCTCCATCAACCTGGGCTTG 4040  
QY 4014 gtttacaaggtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4073



QY 57 gctgcacatggtggggccgtggggcccgccggcgccgagcagcctccgagagc 116  
 Db 79 GCGCCATGCGTGGGGCGCGGGGCGCTCTGGAGCGCGCGCGGCTGAGCCACAGGAGAGC 138  
 QY 117 taccgcacacagagacacagagctacagcgcctggagcccatctacgcgccgacttc 176  
 Db 139 TATTGCGAGCGACAGACACAGAGCTGAGGCGCTGAGGCGCATTTACGGCTCGGACTTC 198  
 QY 177 caagacctgcggccggaagctggtggagccggtcaaaagagcccccctgaatcaattagtt 236  
 Db 199 CAGGATCTGCGCGCGATGCGCGCGAGCGGTGAGAGCGCTCTGTAATCAACTTAGTT 258  
 QY 237 tigtacctcaagcctcaactggtgaaagatatgtaaaaagtgaggttgaggttaaa 296  
 Db 259 CTTTACCCCTCAGGGCTAGCTGGTGAAAGAGTATACGTGCAAGTGAAGTGAAGTGA 318  
 QY 297 tgcacacctaccctatccagatgtagtctcctgaaatagagtttaaaaaatgcaaaagttcta 356  
 Db 319 TGCCCGCCCAACATACCAGATGATGTCCCGAAATAGAGTTAAAAATGCTTAAAGGCGTG 378  
 QY 357 tcaaatgaaagtgtaaatgttgaataatcgcctagagaaactggccaaagaaacactgt 416  
 Db 379 TCGAATGAAAGTGTTAAATGTTAAATCACACCTGGGAAGAACTGGCCAAAGAACAGTGT 438  
 QY 417 gggaggtgatgatcttgaactggctaccacgtgcagtcatttctcagcgagcataac 476  
 Db 439 GGAGAGGTGATGATTTGAACCTGGCACACCATGTCAGTGTCTTCTCAGCGAGCATAAAC 498  
 QY 477 aagccccctcccaagtccttcatgaagaaatgctggaaagcgggctcaggaagagcag 536  
 Db 499 AAGCCCCCAACCAAGTCTTCCATGAAGAAATGCTGGAAAGCAGCGCTCAGGAAGCAG 558  
 QY 537 cagaggtgtgagcccaagcgaaagagagcaggaacggtgaatcctcgtcatg 596  
 Db 559 CAGAGGTTGTGAGGCGCCAGCGGAAGAGAGGAGGAGGACACGTGAATCTCCACGAG 618  
 QY 597 attcagagaaggaaagagataaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 656  
 Db 619 ATTCAGAGAAGGAAGAGAGATCAAGGAAGAGAGAGAAAGGAAGGAAGGAAGGAAGGA 678  
 QY 657 gaaagtgtgaaattgctagtgttcaaaccaagatcatcacttaagaagagaccagga 716  
 Db 679 GAACGTTTGGAAATFCTAGCTTGACAAACCGAGACTATGCTCTTAAGAGAGACCCAGCA 738  
 QY 717 gacacagacggtgctccattctacatggaggtcctcctgaatttgtaggaaatggtaaa 776  
 Db 739 GGGCAGAGGCGAGCTGCCATCCCTCCATGGAGGCTCTCTGACTTTGTAGGAAATGTTAA 798  
 QY 777 catcgggcaaaactcctcaggaaggtctaggcagagagcagtcagtattctgtatagtagt 836  
 Db 799 GCGCGGACATACTCTCAGGAAGTCCAGCGAGAGACGTTCAGTATTCTGTGTAGTGTGT 858  
 QY 837 gaagattctcctgctctgtgaaattctgtatttcaatatggggaggtcctcagtcagctc 896  
 Db 859 GAACCCCTCTCTGCTGTGTGACATCTCCATTTCTCTGTGGGTAGCCCTGATCAGCTC 918  
 QY 897 atggtgcacaaaggaaatgatttggcagtgatgaacaaacttggaaatagttcctaact 956  
 Db 919 ATGTTGCAACAAAGGAGATGTTGGCAGTGATGAGCAGCTGGGAAAGTGGTATACAAT 978  
 QY 957 gcttggaaacagcactggtggt 1016  
 Db 979 GCTTTGGAACACAGCCAGCGGCGAGCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1038  
 QY 1017 aaaaaatgggttccattccttaccagtcacagaaagagaaagagaaatgtaagtgcacaaag 1076  
 Db 1039 ---AGATGGGCCCCCTTGCTTTACCAGTCAAGAGAGAAAGAGAGATTGACAAAGTGC 1095  
 QY 1077 cagattcaagggaacagaaacgaatcactcactggttaaaattgaacccatcaaatgta 1136  
 Db 1096 CAGATTCAAGGAGCAGAAACAGAAATTCAGCTCCCTAGTGAACATGAGCCATCCCAATATC 1155

QY 1137 gtagctaccttgcaatgaatctcaaaagagacagactccatcgtggtgacattta 1196  
 Db 1156 GTCCGCTACTTTGGGATGAACATCCAGAGAGAGGAGGACTCCCAATTGTGATCGACATTCG 1215  
 QY 1197 gtgagacacattagtggtgtctctgtgcacacactgagccactcagggcccaatccct 1256  
 Db 1216 CGAGAGACAGCTCAGTGGCATCTCTCTGCTACACACCTGAGCCACTCGGGCCAGTCCCC 1275  
 QY 1257 gtgcatcagcttcaggtgacacagctcagctcctgtcagggccttgattatctgcacag 1316  
 Db 1276 GCGCCACAGCTGCGCAAGTACACAGACCCAGCTCTGCGAGGCTTGGATTACCTACACAGC 1335  
 QY 1317 aattctgtgtgataaggtcctcagtgcatcctaattcttgggtggatgcagaagggcacc 1376  
 Db 1336 AACTCCGTTGGTGCAACAGTTCTGAGCGCTCCAGTGTCTTGTGTAGATCGCGAGGCGACT 1395  
 QY 1377 gtcaagattacggactatagcatttctaagcgctcgcagacatttcgaaggagagtg 1436  
 Db 1396 GTCAAGATAACAGACTACAGCATCTCTAAGCGTCTGGCAGACATTTGCAAGGAGGATGTA 1455  
 QY 1437 tttagcaaaacccaggttcgttttagtgaacaatcctcctctataaaacggggaagaaa 1496  
 Db 1456 TTTGAGCAAGCTCGAGTTCTGTTTGTAGTGACAGTGCCTTGCCTTATATAACAGAGAA 1515  
 QY 1497 gtagatgttggcgtcttggccttctgctgtcctcagccagcaggaagtggtgga 1556  
 Db 1516 GGGATGTGTGGCTTGGCCTCTCTGCTGTCTCTCAGCAAGGACAGAGAGTGTGG 1575  
 QY 1557 gtagtacctgtgacctcctcagctgacttaccagctgacttccaagattttctaaagaa 1616  
 Db 1576 GAGTATCCGCTGACCATCCCTGAGTGCCTGCGAGCTTCTTCCAAAGACTTCTTGAAGAG 1635  
 QY 1617 tgtgtgtgtgagtgacaaagaaagagtgagtcctccagcaggtgttgaacacacagcttt 1676  
 Db 1636 TGTGTCTGCTGGACGACAAAGAAAGATGGAGCCCTCAGCAGCTGCTGAAACACACAGCTTT 1695  
 QY 1677 ataaatccccagcaaaaaatgctcctagtggaaacaaaagtcctgaagattctcggagga 1736  
 Db 1696 ATAAACCTTCAACCAAGCTGCTTTGTTGGAGCAGAGTCTCTGAAGACTCTGGGGGACAA 1755  
 QY 1737 gattatgtgagactgttattcctagcaacggctaccagtgctgctctctttagtgag 1796  
 Db 1756 GACTACATTGAGACCGTTCATCCCAACACAGCTCCCGCCAGCTCCCGCTTCTTCTCAGTGAG 1815  
 QY 1797 acacagagacagtttcccgatacttcatgttgaagtgtgaagaattacaactcttggtaaa 1856  
 Db 1816 ACACAGAAACAGTTCCTCCGGTACTTCTATTGAGTTTGAAGAACTACACCTCTAGGGA 1875  
 QY 1857 gtagcttttggagctgtcatcaagggtgcagaaacagtttgagcgctgctgctacagctg 1916  
 Db 1876 GGAGCTTTTGGAGCTGTCTATCAAGGTGCAAAACAAAGCTTGATGGCTGCTGCTATGCTGTG 1935  
 QY 1917 aagcgatccccatcaacccggccagccggcagttcccgaggtatcaagggcggaagtga 1976  
 Db 1936 AAGCGATATCTATCAACCTTCAAGCAGACACTTCCGCGGTATCAAGGCGGAGGTGACA 1995  
 QY 1977 ctgctgtcaaggctgacacatgagaaacattgtcgtactataaaccttggatcgagcg 2036  
 Db 1996 CTGCTATCGCGCTGACCATGAGAACATTGTGCGTACTTACAAACGCTTGGATCGAGCGG 2055  
 QY 2037 cagcagcgccggcgagccgggagcccgcccccgaactcccggtcccttggccaaaggt 2096  
 Db 2056 CATGAGCGCCAGCGGGTCCAGGAGACACCGCCCCCAGCTGACACCCAGGCGCCAGGAC 2115  
 QY 2097 gaccgagctgcacgcggggcagcgcgagcagacacagcggcctggacagcgttagagcc 2156  
 Db 2116 AGCCAGCCACTACGGGAAACATCAGGCGACACTGAAGAACTGGGCGAGCTGGAGGCT 2175  
 QY 2157 gccgcgcgcgcacccatcctcagcagctcgttgagtgagcacttcggcgagcgctcg 2216  
 Db 2176 GCAGCAGCGCCACCATCTCTCAGCAGCTCGGTGAGTGGAGACATCTGCAGAGCGTCTCT 2235  
 QY 2217 gccagtgcctgtttcccgccagcccgccgggtcccgagcgtgacgagagcagcag 2276

Db	2236	ACCAAGCACCCTTCCAGCTTCCAGCTACGGCGCAGGATTCAGC---AGCGATAGGAGAGAGCAG	2299
QY	2277	gacgagcagcggctgtcttctccagctcttctgcgtctgcttgcagattctgaaagtgat	2336
Db	2293	GACGAGCGCATGCGCTTCTCCAGTCTCTTACCTGCTCAGATTTCTGACAGTGCAC	2352
QY	2337	attatctttgacaatgaagatgagaacactaaagtcaagaatcaggatgaagattgcgat	2396
Db	2353	ATCATCTTTGACATGGAAGATGAAGAAACAGTAAAGTCAAAATCAGGATGAGACTGCAAT	2412
QY	2397	gaaagaatggctgcatgaaagtggagccatcagtgacgactgagcgtgagcgtgacactaccta	2456
Db	2413	CAAAAGGACGGTAGCCATGAGTTGAGCCTTCAGTGACAGCTGAAGCTGTGCATACCTT	2472
QY	2457	tacatccagatgagctactgtgaagaagacactttacgagacacacttgaccaggagctg	2516
Db	2473	TATATCCAGATGGAGTACTTCGAGAAAGACGCTGCGGGACACCATGACACAGGGCGTG	2532
QY	2517	tatcgagacacgctcagactctggaggctttttcgagagattctggatgattagcttat	2576
Db	2533	TTTCGACACACACCGCGCTCTGGAGGCTTTTCGAGAGATTCGAGATGATTAGCTTAT	2592
QY	2577	atccatgagaagaatgattcacgcgggatttgaagcgtgcacatttttttggattct	2636
Db	2593	ATCCATCGAAGAAAGATGATTCATCGCACTGGAAGCTGTCAATATTTTTTGGATTCT	2652
QY	2637	gatgaccatgtgaaaatagtgatttgggttggcgacagaccatctagcctttctgct	2696
Db	2653	GATGACCATGTAAAAATAGTGTGACTTTGGCCTGGCAACAGACCATCTGGCCTTCACGTCT	2712
QY	2697	gacgacaacagacgatacgaacagaga---cttgattaagtcagacccttcaggtcac	2753
Db	2713	GAAGGTAAACAGACGCGCAGCGAGTGATCAGTGTGATTAAAGTCACACCTTCAGGCCAT	2772
QY	2754	ttaactggatggctggcactgctcttatgaagccagagctccaagaagacacaaa	2813
Db	2773	TTGACTGGCATGTTGGTAGTACTCTGTATGTATGTAAGCCCTGAGGTCCAAAGACACCAAG	2832
QY	2814	tctgcatacaacgaagaagtggatctcttcagcctgggaattattcttttggatgtcc	2873
Db	2833	TCGTGATACACCAAGAAAGTGGATCTCTTCAGCCTGGGAATATATCTCTTTGAGATGTCC	2892
QY	2874	tataccccatggctcacggcttcgaagagatcttgttctcaaccaactcagagatccc	2933
Db	2893	TATCACCCGATGGTGACTGCCTCAGAAAGAAATTTTGTCTTCACCACTTCGAGATCCC	2952
QY	2934	acttgcctaaagtcttcagaagaatttgaagatgagagatgcgaagcaaaatcagtc	2993
Db	2953	ACATGCCCCAGTTTCCAGACGACTTCGATGATGGAGACATACAAAGCAAAATCTCTC	3012
QY	2994	atctcctggctgttgaaocacgataccagcaaaagccccacagccacacagaactgctcaag	3053
Db	3013	ATCTCTGTGCTTTGAACCATGATCCAGCTTAACGGCCCCACGGCCATGGAATGCTCAAG	3072
QY	3054	agtgagctgctgccccacccccagatggaggatcagagctgcatgaagtgtctgcacac	3113
Db	3073	AGTGAGCTGTGTCCCGCCGCGCAGATGGAGGAATCGAGTGCATGAAGTGTCTGCACAC	3132
QY	3114	acgctgaccaacgtggatgggaaggcctaccgacacatgattggccagatcttctcagag	3173
Db	3133	ACGCTGGCAACACGATGGGAAGGGCTACCGACCATGATGAGCGAGCTCTTCTGCGAC	3192
QY	3174	cgactctcccctgcactgatattacactatgacagcgacatactgaagggcaactctca	3233
Db	3193	CACAGCTCCCCCGCATAGACTACACTACGACAGCGACATCCTGAAGGGCAACTTCTTG	3252
QY	3234	atccgtacagccaagatgcagcagcatgtgtgtgaaccacatcatccgcattcttaaaaga	3293
Db	3253	ATTGCGACAGCCAGATCCAGCAGCTGTGTGTGAAACCATCTGTCGCGCTTTCAAAGG	3312
QY	3294	catggagctgtcagttgtgtactccactactgtctcccgaaacagacaaaatatgatgag	3353

Db	3313	CAATGAGAGCTGCCAGCTGTGTGCACCCCGCTGCTCTTCCCGAAACAGGCAATATATGAG	3377
QY	3354	cacaacgaagctgcctattatcagaccacagcgagatcgtatgcttccttcttgac	3413
Db	3373	CACAACGAAGCGCTCTGTTTCATGGATCACAGTGGGATGCTGGTGATGCTTCTTTTCGAC	3432
QY	3414	ctgcggaatcccttttgcgaagatgctggaagaataatattgaaatttaaaacgatac	3473
Db	3433	CTGCGGGTCCCTTTTGCAAGATATGTGGCAAGAAATTAACATACTGAATTTAAAGCGATAC	3492
QY	3474	tgcatagaacgtgtgttcaggccgcgaagtttagatcatcttattcaccagaactctctg	3533
Db	3493	TGCTATAGACGGGTGTTTCAGACCTAGGAAGTTAGACCGGATTTCACTCCCAAGAAGACTCCTG	3552
QY	3534	gagtgatgattgatattgtcaactctacaccaccaacagctttctgccactgctgaaatt	3593
Db	3553	GAATGTGCATTTGACATCGTCACTTCTACCGCCAACAGCTCCCTGCCACTGCTGAAACC	3612
QY	3594	atctacactatctataaaatcatccaaagatttcagacacttcaggaaagaataatacagt	3653
Db	3613	ATCTTACACCATCTATGAAGTCATCCAGGAGTTCTCTGCACCTTCAGGAAAGAAATTAACAGC	3672
QY	3654	atttattgaaccataccattgttattgaaagcaatctctttacactctgggagtcaccagaa	3713
Db	3673	ATTTTACTTGAACCATCAATGTGCTTAAAGGCAATACTCTTACACTCTGGGATCCACAG	3732
QY	3714	gataaacctcagtcgaactacattattctgtatgctgtgacagagaagctgcagagg	3773
Db	3733	GATTAACCTCAGTCAAGCTACGCTTATCTGTATGATGCTGTGACAGAGAAGCTGACTAGG	3792
QY	3774	agagaagtggaaactaaatttgttaactctgtctttgtctcttctaatagtctgtgtcgaatc	3833
Db	3793	AGGGAAGTAGAAGCTAAGTTCTGTAACTGTCTTGTTCATCTAATAGCTGTGCTGTGCACATC	3852
QY	3834	tacaagtttatgaaacagaaggagattgcaagatcttatgcgaacaaataaattcatta	3893
Db	3853	TATTAAGTTTATCGAGCAGAAGGGGATTTGCAAGACCTTAACGGCGACCAATAAACTCACTA	3912
QY	3894	ataaacagaaacacaggtattgtcacagttggctgaagtgtgcttaaaagaccttagagag	3953
Db	3913	ATAAAAACAGAAACAGGCATTTGCTCAGCTGGTGAAGTATAGCTTAAAGACCTTAGAGAA	3972
QY	3954	gttgttgagctgtgaagaactcggcaatcaagttacaggtctctgtatcaatttgggcttg	4013
Db	3973	GTGTGCGGCTGCTGAAGAAACTTTGGCGTGAAGTTGAGGCTTCACCTCAACCTGGGCTTG	4032
QY	4014	gtttacaaggtgcagcagcaaatggaatcatcttcagtttgtgcttctcatcaaacga	4073
Db	4033	GTTTTCAAGGTGCACGACCAATGGCATCATCTTCCAGTTCCTGGCATTCAGCAAAACGC	4092
QY	4074	aggcaaaaggtgtacctgaaatcctcgcagctggaggcagatagactgcgtattcccc	4133
Db	4093	AGGCAGAGGGTTGTGCCCTGAGATGCTTTGCAGCTGGTGGCAGATATGACCTGCTGATTCCC	4152
QY	4134	cagttttagaggccaacagctctggggccagttcccaactgccaactggggctcagcataagct	4193
Db	4153	AAGTTTCAGAGGCCACAGGCTCTGGGGCCAGTCCCCACTGCTGTGGGTGTCAAGCATAGCC	4212
QY	4194	atagacaagatatctgtctgtctcacaatggaggaatctgttacataaagctcttgt	4253
Db	4213	ATAGACAAGATATTTGGCGGCTGTCTTCAACATGGGAGAGCCTGTTTACAGTGAGCTCCTGT	4272
QY	4254	gacctcctgtgttaagtgtgtgcagatgtctatgtccagggccaatcaactaaaccag	4313
Db	4273	GACCTCTGTGTGTGCTGTGGCCAGATGTCATGTCCAGGGCCATCAACCTTAACCCAG	4332
QY	4314	aaactctggacagcaggcatcaacagcagaaatactgacagctggtcacagtcacagtcacag	4373
Db	4333	AAACTCTGGAGCGCGGCATCACTGCAGAGATCATGTATGATCTGCTCAGTCCCGAA	4392
QY	4374	gaattacaagaagtactcaacataatgaatacaactatgtggccctgtctcgtgataa	4433
Db	4393	GAGTTACAGAAATCTGCAGACATCAGAAATCACTATGTGGCCCTGGTCTGTGATAA	4452









QY	2894	ttcagaaagatcttggttctcaacaactcagagatccacccttgcgttaagttccaga	2953	3974	actcgcatcaagttacaggtcttgatcaaatgttggtggttcaaaaggtgcagcaga	4033
Db	3005	CTCAGAAAGAAATTTTGGTTCTCAACCACTTCAGATCCACATCGCCCAAGTTTCCAGA	3064	4085	ACTTTGGCGTGAAGTTGCAAGGTCTCCATCAACTGGCTTGGTTTACAAAGTGCAGCAGA	4144
QY	2954	agactttgacgattggagagatgcaaaagacagaataatcaatcatctctgctgttgaacca	3013	4034	caatggaatcatcttccagtttggcttttcaatacaaacgaagcaaaaggtgctgactga	4093
Db	3065	CGACTTCGATGATGGAGAACATACAAGACAGAAATCTGTCATCTCTCGCTTGTGAAACCA	3124	4145	CAATGGCATCATCTTCCAGTTCCCTGGCATTCAGCAACGACGACGAGAGGTTGTGCTTGA	4204
QY	3014	cgatccagcaaaacgcccacagccacagaaactgctcaagagtgcgtgcgcccccacc	3073	4094	aatctcgcagctggagcgagatagacctgctgattcccccagtttagagggccacaagc	4153
Db	3125	TGATCCAGCTTAACGGGCCACGSCCATGGAACCTGCTCAAGAGTGAGCTGCTCCCCGCC	3184	4205	GATCCTTGCACTGGTGGCAGATATGACCTGCTATTCCCAAGTTTCAGAGGCCACAGGC	4264
QY	3074	ccagatggagagtcagagctgcatgaagtgtgcaacacagctgacacaagctgtagtg	3133	4154	tttggggccagttcccaactgccaattgggggtcagcatagctatacaagaatcatctgctgc	4213
Db	3185	GCAGATGGAGAAATCCGAGCTGATGAAGTGTGACACACGCTGGCCACACAGCATGG	3244	4265	TCTGGGGCCAGTGCACCATGCTGCTGGTGTGAGCATATGCCATATGACCAAGATATTCGCGCG	4324
QY	3134	gaagcctaccgcaaccatgatggccagatcttctcgcagcgcacatctcccctgccaatga	3193	4214	tgctccaacatggaggaatctgttacaaatgaagctctgtgacctcctgggttgtaagtg	4273
Db	3245	GAAGGCTTACCGCACCATGATAGCCAGCTCTTCTGCCAGCAGCTCCCGCGCATAGA	3304	4325	TGTCCTCAACATGGGAGAGCCCTGTTACAGTGAGTCTCTGTACCTCTGTTGTTCAGTGC	4384
QY	3194	ttacacctatgacagcgacatatactgaagggaactctcacaatcgtacagcaaatca	3253	4274	tgttcagatgtctatgttccagggccatcaacctcaaccacagaaactctgacagcagc	4333
Db	3305	CTACACCTTACGACAGCAGCATCTCTGAAGGCAACTTCTTGATTCGCACAGCCCAAGTCCA	3364	4385	TGSCCAGATGTCCATGTCCAGGGCCATCAACCTTAACCCAGAAACTCTCTGGACGGCGGCAT	4444
QY	3254	gcagcatgtgtgaaccatcatccgcacatctttaaagacatggagctgttccagtttg	3313	4334	caacgagaaaatcatgtacgactggtccacagtcccaagaggaattacaagaagtactgcag	4393
Db	3365	CGAGCTTGTGTGAAACCATCTGCCGCTCTTCMAAAGGCATGGAGCTGCCAGCTGTG	3424	4445	CATTCGACAGATCATGTATGACTGGTTCACAGTCCCGAGGAAGTTTACAAGAATACTGCAG	4504
QY	3314	tactccactgcttccccgaacagacaaatatatgagcaacaagagctgcctatt	3373	4394	acatcatgaatacacctatgtggccctgtctcggataaaaaggaagccatgtcaaggt	4453
Db	3425	CACCCCGTGTGCTTCCCGCAACAGSCAAATATATGAGCAACAGCAAGCCGCTCTGTT	3484	4505	ACATCAGAAATCACCTATGTGGCCCTGTCTCTGATTAAGAAAGAAAGCCATGTCAAGGT	4564
QY	3374	caigacacacagcgagctgtgagctctctttgacctgagctgcgttctttgcaag	3433	4454	taagttttcgaaagaagaagcagacagagagtgctgctggagactgaaacttgtgga	4513
Db	3485	CATGGATCACAGTGGCATGCTGGTGTATGCTCTTCGACTGGGGTCCCTTTTCAAG	3544	4565	CAAGTCTCTTTGAGAAGGAGGCAACACAGAAACGTGTGTGGAATCAGATCTTGTGA	4624
QY	3434	atatgtggcaagaaataattgaatttaaaacgatatgcatagaaacgtgtgtcag	3493	4514	ccatgtactgcagaaactgaggactaaagtccactgagtgaaggaatggcagagaagcttc	4573
Db	3545	ATATGTGCAAGAAATAACATACTGAATTTAAAGCGATCTGCTAGAGCGGTGTTCAG	3604	4625	CCATGTTATGCAGAAACTAAGGACCAAGTTGTGTGACGAAAGAAATTTCCAGAGACGCTTC	4684
QY	3494	gcgcgcgaagttagatcgatttcccaagaaacttctgaggtgctgcatgttatgtt	3553	4574	cgaatacttgcagtgcaaaatctgaagggtcatcttctaatgcttcaggtgtgtttga	4633
Db	3605	ACCTAGGAAGTTAGACCGATTTTCATCCCAAGAACTCTCGGAATGTGCATTTGACATCGT	3664	4685	TGATAATCTTGGCGTACAAACTCTGAAGGGTCTGTTTTCTAATGCTTCAGGTTGTTGA	4744
QY	3554	cacttaccacacacagcttctgccactgctgaaattatctacactatctatgaat	3613	4634	aatcatggagcaacagtggtcccatgtgagttgctagccccggagagcgtgcagc	4693
Db	3665	CACCTTACCGCCACAGCTCCTGCCACCTGCTGAACCACTCTACACCATCTATGAAGT	3724	4745	AATCCATGGAACACAGTAGTCCCCACTGTGAGTGTATATACACGAAATACTGTCCGC	4804
QY	3614	catccaagagtttccagcacttcaggaagaataatcaagtatttattgaacctaccat	3673	4694	cagcaactggaggcgctatgaaactcaggtacaaactcaacttcagacctccctgtccaa	4753
Db	3725	CATCCAGAGAGTTTCTGCACTTCAGGAAGAAATTTACAGCAATTTACTTTGAACCATACAAT	3784	4805	CAGCACACGAGGCGACATGAGATTCAGGTGCAAAACCCCGACTTCACACTACCCCTTGCCAA	4864
QY	3674	gttatgaagcaataactcttacactgtgggacccagaagataaactcagtcagtgcta	3733	4754	cttaccatcagaaaagcagtgaaattgaaattctggctgtggttatccccaagaaacaat	4813
Db	3785	GCTCCTAAAGGCAATACTCTACACTGTGGGATCCAGAGGATAAACTCAGTCAAGTCTA	3844	4865	TTTACATCAGAAAGCAGTGAATTTGAAATTTTGGCTGTGGATCTACCCCAAGGAGACCAT	4924
QY	3734	cattattctgtatgactgtgacagagaagctgcagaggagagaagtggagctaaatt	3793	4814	attacagtttttatcatagagtggggatgctgataacagcgcatattaaacaaactgtgaa	4873
Db	3845	CGTATTCTGTATGATGCTGTACAGAGAGTGTACTAGGAGGGANGTAGAAGCTAAGTT	3904	4925	CATTACAGTTTCTACGCTAGAGTGGGACGCTGTATGATGACAGGCATTTTAACACACACCGTGAA	4984
QY	3794	ttgtaactgtcttcttctaagtctgtgcactctcacaagttatttgacaagaa	3853	4874	gcagctgtctcagccctgcgaagcaagataacctcaaatattagctgtgatgaaattta	4933
Db	3905	CTGTAACTGTCTGTATCTAATAGTGTGTGCTGACTCTATAAGTTTCATCGAGAGAA	3964	4985	ACAGCTGTCTCACGCCCTGCCAAAGCAAGATAGCTCAAACTGCTCTGCCACGAAATCTA	5044
QY	3854	gggagatttgcaagatcttatgccaacaataattcattataaacaagaaacaggtat	3913	4934	taacatcaagtagaaaaaagggtgtctgtctattctgtacagctatagagatgacta	4993
Db	3965	GGGGGATTTGCAAGACCTTAACGCCGACCAATAACTCACTTAATAAACAAGAAACAGGCAT	4024	5045	TAACATCAAAAGTTGAGAAGAAGGTGTCGGTGTGTTCTGTCTACAGCTTACAGAGATGACTA	5104
QY	3914	tgcacagtgtgaaagtatggcttaaaagacctagaggaggtgtgtgagctgttgaagaa	3973	4994	ctacagaactcttattttaacccttaagaactgtcgttaacctcattcaaacacagacaggg	5053
Db	4025	TGCTCAGCTGGTGAAGTATAGCTTAAGACNCTAGAGGAAGTTGCGGGCTGCTGAAGAA	4084	5105	CTACAGAACTCT	

Db	588	AGGAACGTTTGGAAATCACTAGCTTGTGACAAACCAGGACTATGCTCTTAAGAGAGACCCAG	647
QY	715	gaggacacagaacgcyctgccattctacatgagagctctcctgactcttagagaaatggta	774
Db	648	CAGGGCACAGGCGAGCTGCCATCCTCCATGAGGCTCTCCTGACTTGTAGGAATGGTA	707
QY	775	aacatgggcgaactctcaagaaggtctaggcgagaaactcagtatctgtatgtaata	834
Db	708	AGCCCGGACATATCTCTTCAGAAAGGTCAGGCGAGAACGTCACTATGTGTGTAGTG	767
QY	835	gtgaagattctctggctctgtgaaattctgtatttcaatagggagagctcctgacagc	894
Db	768	GTGAACCCCTCTCTGGCTCGGTGACATCTCCATTCTCTGTGGGTACCCCTGATCAGC	827
QY	895	tcatgtgtcacaaagggaatgtattggcagtgatgaacaacttggaaaattagttctaca	954
Db	828	TCATGGTGCAAAAGGAGATGTGTGGCAGTGATGAGCAGCTGGGAAAGGTGGTATACA	887
QY	955	atgcttttggaaacagccactgtggctttgttctgttattgtagtgggtcccttcagtgcc	1014
Db	888	ATGCTTTTGGAAACAGCCAGCGGAGCTTTGTCTGTGTCATGAGTGGGTCTTTCAGTGGC	947
QY	1015	agaaaaaatgggtccattctaccagtcagaagaaagaagattgataagtgcaaaa	1074
Db	948	AG--AGATGGGCCCCCTTGCTTTACAGGTCAAGAGAAAGAGAAGATTGACAAGTGCAAAA	1004
QY	1075	agcagattcaaggaaacagaacagaattcaactcactggtaaattgagccatccaaatg	1134
Db	1005	GGCAGATCAAGGAGCAGAAACAGAAATTCAGCTCCCTAGTGAACCTGAGCCATCCAAATA	1064
QY	1135	taqtacactaccttgaaatgaatctcaagagcaagacgaactccatcogtggctgacattt	1194
Db	1065	TCGTCCCGCTACTTTGGCGATGAATCCCAGAGAGAGGAGGAGGACTTCCATTGTATCGACATTC	1124
QY	1195	tagtggagcatttagtggggtctctctgtctgacacctgagccactcaggcccatcc	1254
Db	1125	TGGCAGAGCAGCTCAGTGGCATCTCTCTGGCTTACACACCTGAGCCATCTGGGCCAGTCC	1184
QY	1255	ctgtgataagcttcgaagtcacacagctcagctcctgtcaggcccttgattatctgcaca	1314
Db	1185	CCGCCACACAGCTGGCGAGTACACAGCCAGCTCTCTGGCAGGCTGGATTACTATACACA	1244
QY	1315	gcaattctgtgtgcataaagtcctgagtgcatctaatgtctctgtggatgcagaaggca	1374
Db	1245	GCAACTCCGTGGTGCAAGGTTCTGAGCGGTCCAGTGCTCTGGTAGATGCCGAGGGCA	1304
QY	1375	ccgtcaagattacgactatagcattctcgaagcctcgaagcatttgcgaaggagatg	1434
Db	1305	CTGTCAAGATACAGACTACAGCATCTCTAAGCGCTCTGGCAGACATTTGCAAGAGGATG	1364
QY	1435	tgtttgagcaaacccagttcgttttagtgacaatgctctgccttataaagcggaaga	1494
Db	1365	TATTTGAGCAAGCTCGAGTTCTGTTTGTAGTACATGCCCTGCCCTTATTAACAGGAAGA	1424
QY	1495	aaggagaatttggcgtcttggcctctctgtctgtcctcagccaaggacaggaatgtg	1554
Db	1425	AAGGGATGTGGCGTCTTGGCTCTCTGTGTGTCTCTCAGCCAAGGACAGGAGTGTG	1484
QY	1555	gaggtacccgtgaccatccctagtgaacttaccagctgacttcaagatttcttaaga	1614
Db	1485	GGGAGTATCCGTGACCATCCCCAGTGACCTGCCAGCTGACTTCCAAGACTTCCCTGAAGA	1544
QY	1615	aatgtgtgtgtgatagaagaaagatggagtccccagcagttgttgaacacagct	1674
Db	1545	AGTGTCTTGGCTGGACACAAAGAAAGATGGAGCCCTCAGCAGCTGCTGAAACACAGCT	1604
QY	1675	ttataaatcccccagccaaaaatgctctagtggaaacaaagattcctaaagattctggagac	1734
Db	1605	TTATAAACCCCTCAACCAAGCTGCCTTTGGTGGAGCAGAGTCTCTGAAAGACTCTGGGGAC	1664
QY	1735	aagattatgttgagactgttattcctagaacccggtacccccagtgctgctctcttagt	1794
Db	1665	AAGATCATATGAGACCGCTCATTTCCAGCAACAGCTGCCAGCGCTGGTTCCTACGTG	1724

```

Query Match      61.9%; Score 3418.4; DB 94; Length 5141;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 3835; Conservative 0; Mismatches 611; Indels 10; Gaps 4;
QY 655 aggaacgcttggaaatgctagttgtcaaaccaagatcattacctctaaagaagccacc 714
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```





Query Match		60.3%	Score 3332	DB 93	Length 3351		
Best Local Similarity		100.0%	Prod. No. 0				
Matches 3332		Conservative 0	Mismatches 0	Indels 0	Gaps 0		
QY	2194	ggagcacttcggcgagcgctgcggcagctgcgcgcttccccgcgcacgcggccggctccca	2233				
Db	1	GGAGCACITCGGGCGAGCGCTCGGGCAGTGGCGGTTTCCCGCCACCGGGCCGCTCCA	60				
QY	2254	ggagtag	2313				
Db	61	GGGATGACAGAGACACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	120				
QY	2314	ctgctccagatctcgaagtgatattattctttgacaatgaagatgagacagtagtaaaagtc	2373				
Db	121	CTGCTTCAGATTCTGAAAGTGATATATCTTTGACAATGAAGATGAGAACAGTAAAGTC	180				
QY	2374	agaatcagagtagaagtagaagtagaagtagaagtagaagtagaagtagaagtagaagtaga	2433				
Db	181	AGAATCAGGATGAAGATTGAATGAAGAAATGGGTGCCATGAAGTAGGAGGAGGAGGAGG	240				
QY	2434	cgactgagctgacactac	2493				
Db	241	CGACTGAGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	300				
QY	2494	gagacacattgaccagggactgtatcgagacacacacacacacacacacacacacacacac	2553				
Db	301	GAGACACATTGACAGGGACTGTATCGAGACACACACACACACACACACACACACACACAC	360				
QY	2554	agattcagagtagattagattatccatgagaagaagtagattcaccggagatttgaagc	2613				
Db	361	AGATTCTGGATGATTAGCTTATCCATGAGAAAGAAATGATACCCGGGATTTGAAGC	420				
QY	2614	ctgtcaacatttttggattctctgacacacacacacacacacacacacacacacacacac	2673				
Db	421	CTGTCAACATTTTTTGGATTCTGATGACCATGTGAAATAGGTGATTTTGGTTGGCGA	480				
QY	2674	cagacacattcagcttctctgacagacacacacacacacacacacacacacacacacac	2733				
Db	481	CAGACCATCTAGCTTTCTCTGCTGACAGCAACAAACACACATCAGACAGGAGACTTGATTA	540				
QY	2734	agtcagaccttcaggtcacttaactggagaggttggcactgctctctatgtaagccag	2793				
Db	541	AGTCAGACCTTTCAGGTCACTTAACCTGGGATGTTGGCATCTCTATGTAAGCCAG	600				
QY	2794	aggtccagaagaagcaccac	2853				
Db	601	AGGTCCAGGAAGCACCAATCTGCATACAAACACAGAAAGTGGATCTCTTCAGGCTGGGA	660				
QY	2854	ttatctcttggagatgctcctacacacacacacacacacacacacacacacacacacac	2913				
Db	661	TTATCTCTTTTGGATGTCTTATCACCCCATGGTTCAGGGCTTCAGAAAGATCTTTGTTT	720				
QY	2914	tcaaccaactcagagatcccaacttcgcttaagtttccagaagaactttgacagatggagagc	2973				
Db	721	TCACCAACTCAGAGATCCCACTTCGCTTAAGTTTCAGAAAGACTTTTGACGATGGAGAGC	780				
QY	2974	atgcaagcagaaatcagtcactcctcctggctgttgaaccacacacacacacacacacac	3033				
Db	781	ATGCAAGCAAAATCAGTCTCTCTGGCTTGTGAACACGATCCAGCAAAAGGGCCCA	840				
QY	3034	cagccacagaactgctcaagatgagctgctcccccacacacacacacacacacacacac	3093				
Db	841	CAGCCACAGAACTGCTCAAGAGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900				
QY	3094	tgcataagagtgctgcac	3153				
Db	901	TGCTGAAGTGTGCAACACAGCTGACCAAGCTGGATGGGAGGCTTACCGCACCATGA	960				
QY	3154	tggccacagattcttcgagcgacacacacacacacacacacacacacacacacacacac	3213				
Db	961	TGGCCACAGATCTTCTCGAGGCGATCTCCCTGGCATGCTTACACCTTATGACAGCGACA	1020				

QY	3214	tactgaaggcgaaacttctcaatccgtacagccaagatcagcagcatgtgtgtgaacca	3273				
Db	1021	TACTGAAGGCGAACTTCTCAATCCGTACAGCCAAGATGAGCAGCATGTGTGTGAACCA	1080				
QY	3274	tcacccgcatctttaaagacatggagctgttcagttgttactccactactgtctcccc	3333				
Db	1081	TCATCCGCATCTTTAAAGACATGGAGCTTTCAGTTGTACTCCACTACTCTTCCCC	1140				
QY	3334	gaacagacaaaaatatagacacaaagagctccctattcatggaccacagcgagatgc	3393				
Db	1141	GAACACACAAATATATGACACAAAGCTGCCCTATTTCATGGACACAGCGGATGC	1200				
QY	3394	tgtgtgcttccctttgacctgcgagatcccttttgaagatatgtggcaagaataata	3453				
Db	1201	TGGTGAATGCTTCTTTTACCTGGGATCCCTTTTGCAGATATGTGGCAAGAAATAA	1260				
QY	3454	tattgaatttaaaccgatactgcatagaaacgtgttcaggccgcgcaagttagatcat	3513				
Db	1261	TATTGAATTTAAACGATAGTGCATAGAACGTGTGTTTCAGGCGCGCAAGTTAGATCGAT	1320				
QY	3514	ttcatccaaagaactctgagtgatgctcatttgatattgtcaactctaccaccaacagct	3573				
Db	1321	TTTCATCCCAAGAACTTCTGGAGTGTGCATTTTGATATTTGCTACTTCTTACCAACACAGCT	1380				
QY	3574	tctgcccactgctgaatttatctacactatctatgaataatccaaagagtttccagcac	3633				
Db	1381	TTTCGCCACCTGCTGAAATTTATCTACACTATCTATGAAATCATCCAAGAGTTTCCAGCAC	1440				
QY	3634	ttcaggaaagaataatacagatttatttgaaacataccatgttattgaagcaataactct	3693				
Db	1441	TTTCAGGAAAGAAATTTACAGTATTATTGAAACCATACCATGTTTATTGAAACCAATACTCT	1500				
QY	3694	tacactgtggatcccaagaataaaactcagtcagtcacatttctctgtatgatgtg	3753				
Db	1501	TACACTGTGGGATCCCAAGACATAAATCACTCAGTCAAGTCTACATTTTCTGTATGATGCTG	1560				
QY	3754	tgacagagaagctgacgaggagagaagtggaaagctaaatttggtaactgtcttctgtctt	3813				
Db	1561	TGACAGAGAAGCTGACGAGGAGAGAGTGAAGCTTAAATTTTGTAAATCTGCTTTTGTCTT	1620				
QY	3814	ctaatagctgtgtgactctacagatttattgaacagaaggagatttgcagaatctta	3873				
Db	1621	CTAATAGTCTGTGTGCTGCTCTACAAAGTTTATGACAGAGGAGATTTGCAAGATCTTA	1680				
QY	3874	tgccacaataattcattataaaacagaaacaggttattgcacagttgtggaagtatg	3933				
Db	1681	TGCCAACATAAATTCATTTAATAAACAGAAACAGGATTTGCACAGTTGTTGAAGTATG	1740				
QY	3934	gcttaaaagacccatagagaggtgtgtgactgttgaagaacactcgcatcaagttacagg	3993				
Db	1741	GCTTAAAGACCTTAGAGGAGGTTGTTGGACTGTTGAAAGAACTCGGCATCAAGTTACAGG	1800				
QY	3994	tcttgatcaatttggctgtgttcaaggtgcagcagcacaatggaatcatctccactg	4053				
Db	1801	TCTTGATCAATTTGGCTGTTGGTTTACAGGTGCAGCAGCACAAATGGAATCATCTTCCAGT	1860				
QY	4054	ttgtgcttctcatcaaacgaagcaagggctgtacctgaatccctcgagctggagga	4113				
Db	1861	TTGTGCTTTTCATCAAAACGAAGCAAGGGCTGTACTGAAATCTCTCGCAGCTGGAGGCA	1920				
QY	4114	gatatacctgtgattccccagtttagaggccacaagctctggggcagttccccactg	4173				
Db	1921	GATATGACCTGCTGATTTCCCAAGTTTAGAGGGCCACAGCTCTGGGGCCAGTTCCCACTG	1980				
QY	4174	ccatttgggtcagcatagctatagacaagatactctgctgtctgtcctcaacatggagaaat	4233				
Db	1981	CCATTTGGGTCAGCATAGCTATAGACAAGATACTGCTGCTGCTCTCAACATGGAGGAAT	2040				
QY	4234	ctgttacaataagactctgtgacactcctcctcctcctcctcctcctcctcctcctcctc	4293				
Db	2041	CTGTTACAATAAGCTCTTGTGACCTCTCTGCTTGTAAAGTGTGTGTCAGATGCTATGTCCA	2100				



QY	4294	gggccaataacctaaccacgaataactctggacagcagcgatcacagcagaatacatgtacg	4353
Db	2101	GGGCATCAACCTAACCCAGAAACTCTGGACAGCAGGCATCACACAGAAATCATGTACG	2160
QY	4354	actggtcacagttcccaagaggaaattacaagagtactgcagacatcatgaatacacctatg	4413
Db	2161	ACTGGTCACTAGTCCCAAGAGGAATTACAAGAGTACTGCAGACATCATGAAATCACCTATG	2220
QY	4414	tggcccttgtctcggataaagaaggagccatgtcaaggttaagtcttttcgagaagaaa	4473
Db	2221	TGGCCCTTGTCTCGGATAAAGAGGAGCCATGTCAAGGTTAAGTCTTTCGAGAAGGAAA	2280
QY	4474	ggcagacagagaagcgtgtcgtggagactgaacttgtggaccatgtactgcagaaactga	4533
Db	2281	GGCAGACAGAGAAGCGTGTGCTGGAGACTGAACTTGTGGACCATGTACTGCAGAAACTGA	2340
QY	4534	ggactaaagtccactgatgaagaagatggcagagaagcttcgataaactctgcagtgcaaa	4593
Db	2341	GGACTAAAGTCACTGATGAAGAGATGGCAGAGAAGCTTCCGATAATCTTGCAGTGCAAA	2400
QY	4594	atctgaaggggctatttcttaatgcttcaggttttgttgaataccatggagcaacagtgg	4653
Db	2401	ATCTGAAGGGGTCAATTTCTAATGTCTCAGGTTGTGTAATCCATGGAGCAACAGTGC	2460
QY	4654	ttcccatctgtgagtgctagcccgagagaagctgcagccagcaactagagagcgctatg	4713
Db	2461	TTCCCATTTGTGAGTGTGTAGCCCCGGAGAAGCTGTCCAGCCAGCATTAGAGAGCGCTATG	2520
QY	4714	aaactcaggtcaaaactcgacttcagacctcccttgcgaacttcatacagaaagcagtg	4773
Db	2521	AAACTCAGGTCAAACTCGACTTTCAGAGCTCCCTTGCCAACTTACATCAGAAAGCAGTG	2580
QY	4774	aaattgaaattctgctgtgatatcccaagaagaaacaattattcacagtttttatcatag	4833
Db	2581	AAATTGAAATTTCTGCTGTGGATCTACCCAAAGAAACAATATTACAGTTTTTATCATTAG	2640
QY	4834	agtgggagtctgatgaacaggcatttaacacaactgtgaagcagctgcgtcacgcctgc	4893
Db	2641	AGTGGGATGCTGATGAACAGGCATTTTAACACAACCTGTGAAGCAGCTGCTGCACGCCCTGC	2700
QY	4894	caaaccaagatcacctcaaatagctgtgatgaattttaacatacaagttagaaaaaa	4953
Db	2701	CAAAGCAAGATACCTCAAAATTAGCTGTGATGAAATTTATAACATCAAAAGTAGAAAAA	2760
QY	4954	agggtctgtgctatttctgtacagctatagatgactactacagaactcttattttaac	5013
Db	2761	AGGTGCTGTGCTATTTCTGTACAGCTATAGAGATGACTACTACAGAACTTATTATTTAAC	2820
QY	5014	cttaagacactgtgttaacctcaataacacagacagagcgttatactggaaatagtaa	5073
Db	2821	CCTAAAGAACTGTGTTAACCTCAATCAACACAGACAGAGCTTATACTGGAATAATGGAA	2880
QY	5074	tgttgtacatcatcaaatttaaaattaaattcttaagaagagcgtgggtgcagtggtc	5133
Db	2881	TGTTGTACATTCATCAATAATTTAAATTTAAATTTCAAGAAGAGGCTGGGTGCAGTGGCTC	2940
QY	5134	acaccttaactccacagcaatttgggaagccaagggcaggagactgcttgaacacaggagt	5193
Db	2941	ACACCTTTAATCCCAGCACCTTTGGGAAGCCAAAGGCAAGGAGACTGCTTTGAAACACGAGT	3000
QY	5194	ttgagaccgctgagcaacaagcgaagcccccattctctataaaactataaaataatagt	5253
Db	3001	TTGAGACAGCGCTGAGCAACAAAGCAAGACCCCACTCTATATAAAACTTAAAAAATTAGT	3060
QY	5254	tggcataggtggacatgctgtagttccagctactccagagcgtgagatggatcatctg	5313
Db	3061	TGGCATTGGTGACATGTCCTGTAGTCCAGCTACTCCAGAGGCTGAGATGGATCATCTG	3120
QY	5314	agcctcaggaggttgagggtcgagtgagctgtgactgagccactgcactcagctcgtgga	5373
Db	3121	AGCCTTCAGAGGTTGAGGCTGCAGTGAGCTGTGACTGCGCCACTGCACCTCAGTCTCGGA	3180
QY	5374	caacagagcaagaccctgtcttaaaaaaaagaaaaaaatttttttctaagaagc	5433

Db	3181	CAACAGAGCAACGCTGCTCTAAACAAAAAGAAAAAATTTTTTCTAAGAAGC	3240
QY	5434	tgtcctcaaaagtgtgagctttagtttagttttcatgtgtgaatatattataattatcttt	5493
Db	3241	TGTCCTACAAAGTTGAGCTTGTAGTTTTCATGCTGTAATATATTAATTTATCTTT	3300
QY	5494	tgggataataaagtctgtttcatatcacctgca	5525
Db	3301	TGGGATAATAAATGCTTTCATACCTGCA	3332
RESULT	8		
LOCUS	AK027011	2986 bp	29-SEP-2000
DEFINITION	AK027011	2986 bp	29-SEP-2000
ACCESSION	AK027011	2986 bp	29-SEP-2000
VERSION	AK027011.1	GI:10440017	
KEYWORDS	AK027011.1	GI:10440017	
SOURCE	AK027011.1	GI:10440017	
ORGANISM	AK027011.1	GI:10440017	
REFERENCE	AK027011.1	GI:10440017	
AUTHORS	AK027011.1	GI:10440017	
TITLE	AK027011.1	GI:10440017	
JOURNAL	AK027011.1	GI:10440017	
REFERENCE	AK027011.1	GI:10440017	
AUTHORS	AK027011.1	GI:10440017	
TITLE	AK027011.1	GI:10440017	
JOURNAL	AK027011.1	GI:10440017	
COMMENT	AK027011.1	GI:10440017	

**FEATURES**  
**source**

```

agency).
URES
source
Location/Qualifiers
1. .2986
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone="HEP14996"
/clone_lib="HEP"
/note="cloning vector pME18SFL3"
40. .2460
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15625.1"
/db_xref="GI:10440016"

```

## CDS

CDS  
40 2460  
/note="cloning vector pME18SFL3"

```
/count_start=1  
/protein_id="BAB15625.1"  
/db_xref="GI:10440018"
```

/translation - MIHRODOLPWNIFELDDHHVKIGDFGLATDHLAFSAQSKQDQGT  
 DLINSDPSGHLTGMVGTALYSVEVQSTQSKYAINQKDFVLSGIFIFEMSVHPWVLT  
 ERIFVNLQDZSPKPFDDGEHAKQSVISWLNHDLAKPRTATELLKSELLP  
 PQMESSELHILHTINVDGKARVPMQAIFQSRISADIVTDSDDLKGNFSIRTA  
 KMOQGHVEITVIRFKHGAVOLCTPCLLPNRSQIYEHNEALFMDHSGMLVLPFL  
 IFFARYVARNILNLCIKCYERIPRKLDRHPHKELECFADIVTSTTNSFPTAEI  
 IYTVIEIQEAPLQERNYSILNHTMILKAILHCGTPEBKLSQVYIILDAVTKL  
 TREVEAKPCNLSSNSLCRYKFTQKQDQDLIMPSTKQVYIILDAVTKL  
 DLEEVGLPCNLISLNSLCRYKFTQKQDQDLIMPSTKQVYIILDAVTKL  
 YDLIPQFGKQALGPVTAIGVYVQVQHNGIFQFQVAKRRQRAVPEILAQMS  
 SRAINLTOKWTAGTAEITWMSQSQEELQGRHBEITVYALVDSKGSHVYKAF  
 EKERQETRVLETHVQMLQKRTVTDNRNGREASDNLAVNLKGSFNSAGLPEI



HGATVPIVSVLAPEKLSASTRRRYETQVOTRLQTSIANLHKRSSEIEILAVDLPKET ILOFLSLEWDADEQAFNTTVKQLLSRLPKORYLKLVCDEIYNIKVEKKVSVLPYSYR DDYYRIIF"									
BASE COUNT	921 a	641 c	663 g	761 t					
ORIGIN									
Query Match	53.4%; Score 2949.6; DB 89; Length 2986;								
Best Local Similarity	99.8%; Pred. No. 0;								
Matches 2963; Conservative	0; Mismatches 4; Indels 1; Gaps 1;								
y	2553	gagattcgtatggattagcttatccatgagaagaagattcaccgggatttgaag	2612						
b	1	GAGATCTGGATGGATTAGCTTATATCCATGAGAAGGAATGATTCACCGGGATTGAAG	60						
y	2613	ctgtcaacatttttttgattctctgataccaatgtgaataagtgatttggttggcg	2672						
b	61	CTGTCAACATTTTGTGATTTTGAATTTGATGACCATGTGAAAATAGGTGATTTGGTTGGCG	120						
y	2673	acagaccatctagccttttctgtacagcaacaagaagcagatcacagaggagacttgatt	2732						
b	121	ACAGACCATCTAGCCCTTTCTGCTGACAGCAACAAGACGATCAGACAGGAGACTTGATT	180						
y	2733	aadtcagaccctcaggtcacttaactggatgggttggcaactgctctctatgaagccca	2792						
b	181	AAGTCAGACCCCTTCAAGTCACITTAAGTGGATGGTTGGCACATGCTCTCTATGTAAGCCCA	240						
y	2793	qaggtccaggaagcaccataatgcatacaaccagaaagtggatctctctcagcctggga	2852						
b	241	GAGGTCAGAGGAAGCACCATAATCTGCATACACAGAAAGTGGATCTCTTACGCTGGGA	300						
y	2853	attatctcttttgatgtctctatcaccccatgggtcagcgggttcagaagagatcttgtt	2912						
b	301	ATTATCTTCTTTGAGATGTCTATACCCCATGGTCACGGCTTCAGAAAGGATCTTTGTT	360						
y	2913	ctcaaccaactcagagatccactcgcctaaagttccagaagactttgacgatggagag	2972						
b	361	CTCAACCAACTCAGAGATCCCACTTGCCTTAAGTTTCCAGAGACATTTGACGATGGAGAG	420						
y	2973	catgcaagcagaaaaatcagtcactcctcgtgcttgaaccaagatccagcaaaacggccc	3032						
b	421	CATGCAAGCAGAAATCAGTCATCTCTGCTGCTTTGAACCCACGATCCAGCAAAACGGCCC	480						
y	3033	acagccacagaactctcaagagtgcgtcgtccccccacccacagatggagaggtcagag	3092						
b	481	ACAGCCACAGAATCTGCTCAAGATGAGCTGCTGCCCCACCCACAGATGGAGGAGTCAGAG	540						
y	3093	ctgcataagtgctgcaccacacgctgaaccaagctggatgggaagggcctaccgcaccatg	3152						
b	541	CTGCATGAAGTGCTGCACCACGCTGACCAACGTGGATGGGAAGGCCTACCGCACCATG	600						
y	3153	atggccagatctcttcgcagcgcacatccctcgtccatgattacacctatgacagcgac	3212						
b	601	ATGGCCAGATCTTCTCGCAGCGCATCTCCCTGCCATCGATTACACCTATGACAGCGAC	660						
y	3213	atactgaaggcgaacttctcaactcgttacagccgaagatgcagcagcagtggtgtgaaacc	3272						
b	661	ATATCTGAAGGGCAACTTCTCAATCCGTACACCCCAAGATGCAGCAGCATGTGTGTGAACCC	720						
y	3273	atcatccgcacattttaaagacatggagctgttcagttgtgtactccactactgcttccc	3332						
b	721	ATCATCCGCATCTTTAAAGACATGGAGCTGTTTCAGTTGTGTACTCCACTACTGCTTCCC	780						
y	3333	cgaaacagacaaaatatatgacacaacaagaactgcccctattcatcaggaccacagcggatg	3392						
b	781	CGAAACAGACAAATATATGAGCACAACCAAGCTGCCCTATTTCATGGACCACAGCGGGATG	840						
y	3393	ctgtgtgactctcttttgacctcggatcccttttgcaagatatgtggcaagaaataat	3452						
b	841	CTGGGTGATGCTTCTTTTGACCTCGGGATCCCTTTTGGCAAGATATGTGGCAAGAAATAAT	900						
y	3453	atatattgaatttaaaacgatactgcatagaacggtgtgttcaggccgcgcgaagttagatoga	3512						

Db	901	ATATTGAATTTAAACGATACTGCATAGAACGCTGTGTTCCAGCGCGCAAGTTAGATCGA	960						
QY	3513	tttcatcccaaaagaactcttgagtgatgatttgatgatatttgatcaactctaccaccaacacg	3572						
Db	961	TTTCATCCCAAGAACTTCTGGAGTGTGCATTGTGATTTGATTTGTCTACCCACCAACAGC	1020						
QY	3573	tttctgcccactctgaaattatctcacactatctatgaataatcatccaagagtttccagca	3632						
Db	1021	TTTCTGCCCACTGCTGAATATNTCTACACTATCTATNGAATCATCCCAAGAGTTTCCAGCA	1080						
QY	3633	cttcagaaagaattacaglattattttgaaccataccatgtatttgaagaagaataactc	3692						
Db	1081	CTTCAGAAAGAAATTCACATATTATTTGAACCATACCATGTTATTTGAAGCAATACTC	1140						
QY	3693	ttcacatgtgggattcccagaagataaactcagccaagctctacattattctgtgtatgact	3752						
Db	1141	TTACACTGTGGGATCCCAGAGATAAACTCAGTCAAGCTACATATTCTGTATGATGCT	1200						
QY	3753	gtgacagagaagctgacgagagagaagtgaagctaaatttttgaatactctgtctctgtct	3812						
Db	1201	GTGACAGAGAAGCTGACGAGAGAGAAGTGAAGCTAAATTTTGAATCTGTCTTTGTCT	1260						
QY	3813	tctaatagctgtgtcgaactctacaaagtatttgaacagaaggagagatttgcagaactctt	3872						
Db	1261	TCTAATAGTCTGTCTGACTCTACAAGTTTTATTGAACAGAAGGAGATTTGCAAGATCTT	1320						
QY	3873	atgccaacaataaattcattataaacaacagaaacaggtattgcacagtttggtgaagtat	3932						
Db	1321	ATGCCAACAAATAAATTCATTAATAAACAGAAACAGGTATTGACAGTTGSGTCAAGTAT	1380						
QY	3933	ggcttaaaagcctagagaggtgtgttgactgttgaagaactcgcgcagctcagctacag	3992						
Db	1381	GGCTTAAAGACCTAGAGGAGGTTGTTGACTGTTGAAGAAACTCGGCATCAAGTTACAG	1440						
QY	3993	gtcttgatcaaatgttggttttacaagtgacagagcacaatgaatcatctctccag	4052						
Db	1441	GTCTTGTATCAATTTTGGGCTTGGTTTACAAAGTGTGACAGCAACATGGAAATCATCTCCAG	1500						
QY	4053	tttgtgctttctcaaacgaaggaagggcgtgtacgtgaatacctcgcgcagctggagcc	4112						
Db	1501	TTTGTGGCTTTTCATCAACGAGGCAAGGGCTGTACCTGAAATCCTCGCAGCTGGAGCC	1560						
QY	4113	agatatgaactgtctgattcccccagtttagagggccacaagctcttgggccagttcccaat	4172						
Db	1561	AGATATGACTGCTGATTTCCCAAGTTTAGAGGCCACAAAGCTCTGGGCCAGTTCCCACT	1620						
QY	4173	gcatattgggtcagcatagctatagacaagatatctgctgctcctcaacaatgagga	4232						
Db	1621	GCCATTGGGGTCAGCATAGCTATAGACAAGATATCTGCTGCTCTCCTCAACATGGAGAA	1680						
QY	4233	tctgtttacaataagctcttgacccctcctgggttgtaagtgttggtcagatgctatgtcc	4292						
Db	1681	TCGTGTTACAATAAGCTCTTGTGACCTCTCTGTTGAAGTGTGGCCAGATGCTATGTCC	1740						
QY	4293	agggccatcaacctaaaccagaaactcttgacagagcagcatcacagcagaaatcatgtac	4352						
Db	1741	AGGGCCATCAACCTTAACCCAGAAACTCTGGACAGAGGATCATCACACCAAAATCATGTAC	1800						
QY	4353	gactggtcacagtcccaagagggaattacaagagtagctgcagacatcatgaaatcaacctat	4412						
Db	1801	GACTGTACAGTCCCAGAGGAATTTACAAGAGTACTGTCAGACATCATGAAATCACCTAT	1860						
QY	4413	gtggccctgtctcggataaagaaggaagccatgtcaaggttaagtcttttcgagaagga	4472						
Db	1861	GTGGCCCTTGTCTCGGATAAAGAAAGGCAATGTCAAGSTTAAGTCTTTCGAGAAGGAA	1920						
QY	4473	agcagacagagaagcgtgtgtgagactgaactgtgacctgtactgtcagaaactg	4532						
Db	1921	AGGCAGACAGAGAAGCGTGTGCTGGAGACTGAACTTGTGGACCATGTACTGCAAGAACTG	1980						
QY	4533	aggactaaactcactgatgaaaggaatggcagagaagcttcccagataatcttgcagtgcaa	4592						
Db	1981	AGGACTAAAGTCACTGATGAAGGAATGGCAGAGAGCTTCCGATATATCTTGCAGTGCA	2040						

QY	4593	aatctgaaggagggtcatcttcttaa	gtcttcaagggtttgttgaaatccatggagcaacagt	4652
Db	2041	AATCTGAAGGGGTCA	TTTCTTAATGCTTCAGGTTGTTTGAATCCATGGAGCAACAGTG	2100
QY	4653	gttcccatgttgagtgtgctagccc	ccgaggaagctgtcagccagcactcaggaagcgcctat	4712
Db	2101	GTTCCATGTGTGAGTGTCTAG	CCGCCGGAAGAGCTGTGAGCCAGCACCTTAGGAGGCGCTAT	2160
QY	4713	gaactcagggtacaaaactcagact	cagactcccttgcgaacttacctacagaaaaacagt	4772
Db	2161	GAAACTCAGGTACAACTCGACT	TCAGACTCCCTTGCCACTTTACATCAGAAAGCAGT	2220
QY	4773	gaattgaattcttgctgtgtggtat	ctaccocaaagaaaaaatattacagttttttatcat	4832
Db	2221	GAATTTGAAATTCCTGGCTGTG	GGATCTACCCAAAGAAACAATATTACAGTTTTATCATTA	2280
QY	4833	gagtgggatgctgatgaacaggc	attacacaaactgtgaagcagctgcctgcacgcctg	4892
Db	2281	GAGTGGGATGCTGATGAACAG	GCATTTACACACACTGTGAAGCAGCTGCTGTCCAGCCTG	2340
QY	4893	ccaaagcaagagataacctcca	aaattagctctgtgatgaatttataacatcaaatgagaaaa	4952
Db	2341	CCAAAGCAAGATACCTCAAA	TAGTCTGTGATGAATTTATAACATCAAGTAGAAAA	2400
QY	4953	aagtgctctgtctattctgtac	agctatagagatgactactacagaattcttatittaa	5012
Db	2401	AAGTGCTGTGCTATTCTGT	TACACTATAGAGATGACTACTACAGAACTCTTATTTTAA	2460
QY	5013	ccctaaagaactctgtgtta	acctctatccaaacagacagaggtcttactggaataatgga	5072
Db	2461	CCCTGAAGAAGCTGTCAAT	AACCTCAATTCAACACAGACAGAGCTTATACTGGAATAATGGA	2520
QY	5073	atgtgtacattcatcataatt	taaaattaaattctaagaaagctgggtgcagtcgct	5132
Db	2521	ATGTTGTACATTCATCA	TAAATTAATTTCTAAGAAGAGGCTGGGTGCAGTGCT	2580
QY	5133	cacaaccttaatcccagc	actttgggaagccaaagcggaagactcgttgaaaccaggag	5192
Db	2581	CACACCTTTTAATCCCAG	CACTTTGGGAAGCCAAAGCAGGAAGACTGCTTGAAACCCAGGAG	2640
QY	5193	tttgagaccagcctgagca	aaacaaagcagaccccatctctataaaactaaaaaaattag	5252
Db	2641	TTTGAGACAGCCTGAGCA	CAACAGCAGACCCCATCTCTATAAAACCTAAAAAAATTAG	2700
QY	5253	ttgggcatggtggacatgc	gttagtccagctactccagaggtcgagatggatcatct	5312
Db	2701	TTGGGCATGTTGGACAT	GCCTGTAGTCCCAGCTACTTCCAGAGGCTGAGATGATCATCT	2760
QY	5313	gagcctcaggaggttgagg	ctgcagctgagctgactggccaactgcactccagctcgtgg	5372
Db	2761	GAGCCTCAGGAGTTGAG	CTGCTGAGCTGTGACTGTGCTGAGCTGCACCTGCACCTCCAGTCTGGG	2820
QY	5373	acaacagagcaagccctgt	ctt-aaaaaaaaaagaaaaaaattttttttctaaaga	5431
Db	2821	ACAACAGACAGACCCCT	GCTTTAAAAAAGAAAAAAGAAAAAATTTTTTCTTAAGAA	2880
QY	5432	gctgctctacaagttgag	cttgttagtttttcactgtgtaatatattataaattatct	5491
Db	2881	GCTGTCTTACAAGTTGAG	CTTTGTTGTTTTTCATGTGTAATATTATAAAATTTATCT	2940
QY	5492	tttgggatataaatgctt	ctcatata 5519	
Db	2941	TTTGGGATATAATAAT	GCTTTCATATA 2968	
RESULT	9			
LOCUS	HS243428	2162 bp	mRNA	PRI
DEFINITION	HS243428	Homo sapiens partial mRNA for putative eIF2 alpha kinase (GCM2 gene).		15-OCT-1999
ACCESSION	AJ243428			
VERSION	AJ243428.1	GI:6065913		

KEYWORDS	eIF2 alpha kinase; GCN2 gene.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 2162)
TITLE	Berlang, J.J., Santoyo, J. and De Haro, C. Characterization of a mammalian homolog of the GCN2 eukaryotic initiation factor 2alpha kinase Eur. J. Biochem. 265 (2), 754-762 (1999)
JOURNAL	99435990
MEDLINE	2 (bases 1 to 2162)
REFERENCE	Santoyo, J.
AUTHORS	Direct Submission
TITLE	Submitted (28-JUN-1999) Santoyo J., Molecular Biology, Centro de Biologia Molecular "Severo Ochoa", Fac. Ciencias. U.A.M., Cantoblanco, Madrid, 28049, SPAIN
JOURNAL	
FEATURES	Location/Qualifiers
source	1..2162
gene	/organism="Homo sapiens" /db_xref="taxon:9606"
CDS	1..2144 /gene="GCN2"  <1..1649 /gene="GCN2" /function="protein synthesis control" /codon_start=3 /product="putative eIF2 alpha kinase" /protein_id="CAB58360.1" /db_xref="GI:6065914"
	/translation="ALFMDSGLVMLPDLRIPARVARNNILIKRYCIERVPFP MLKDRFPIKELLCADPDIQSVYIILNDYATEKTREVEAKFCNLSSLSICRLIKFI MKLAILHGCSIEPKDLSOVYIILDYATEKTREVEAKFCNLSSLSICRLIKFI EQKGDLQDLPTINSLIKOTGIAQLVTLGLKDLEEVGLLKIGIKQLVLINGLVI IQKHNGIIQVFATKKRKORAVPEILAAGRYDLLIPOFRPOALGPVPTAIGVSIA IDKLSAIVLMNESVTISCDLLVSVSQMSRNLIQTQMTAGITABIIMYDWOS OETDEQRCHHEITYVALVDKRGSHVKFKPEROTKRVLETUVDHLVLOKLRTK VELDNREGASDLNAVKNLGSGFNAGLFTEHGATVPIVSVLAPEKLSASTPRRYE TOVORTLTSLANHOKSEIELAVDLPKETILQFLTSLEWDADEQNFNTYKQLLSLR LPKORYLUKLSDNIYINIKVEKSVLVLYSYRDYYRIIF"
	<2139..2144 polya_signal /gene="GCN2"
BASE COUNT	672 a 434 c 481 g 575 t
ORIGIN	
Query Match	39.1%; Score 2160.4; DB 93; Length 2162;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2161; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	3364 ctgacctattcatggaccacagcgggatcgctgtgatgcctctttgaacctcgcgatcc 3423
Db	1 CTGCCCTATTTCATGGACACACGGCGGATCGTGTGATGCTTCCTTTTGACCTGGGATCC 60
QY	3424 cttttagaatatgtggcaagaataataattgaatttaaacacgactgcataaac 3483
Db	61 CTTTTGCAGATATGTGGCAAGAAATAATATATTGATTTTAAACGACTACTGCATGAAC 120
QY	3484 gtgtgttcaggccgcgaagttagatcgatttatccccaagaactcttgagtgatcat 3543
Db	121 GTGTGTTTCAGSCCGCGCAAGTTAGATCGATTTTCATCCCAGAAGACTTCTGGAGTGTGCAT 180
QY	3544 ttgatatgtcaacttctaccacaacagctttctgccactgctgaaattatctacacta 3603
Db	181 TTGATATTTGTACTTCTACCACCAACAGCTTTCTCCCCACTGCTGAAATTATCTACACTA 240
QY	3604 tctatgaataatcccaagagtttcagacacttcaggaaagaataatcacgattatttta 3663
Db	241 TCTATGAATCATCCAGAGTTTCAGACATTCAGAACGAAGAAATACAGTATTATTTTGA 300
QY	3664 accatacatgttatgaaagaactacttaccactgtgggattccccagaagataaaccta 3723
Db	301 ACCATACCATGTTATTGAAGCAATCTCTACACTGTGGATCCCAAGAATAAACTCA 360

QY 3724 gtcaagctacattatctgtatgtgtgacagagaagctgacgagggagagtg 3783  
|||||  
Db 361 CTCAGCTACATTATCTGTATGCTGTGACAGAGAAGCTGACGAGGAGAGTGG 420  
|||||  
QY 3784 aagctaaatttgaatctgtcttctgtcttcttaataagctgtgtgactctacaagtta 3843  
|||||  
Db 421 AAGCTAAATTTGTAATCTGTCTTTGTCTCTTAATAGTCTGTGCTCAAGTTTA 480  
|||||  
QY 3844 ttgaacagagagatttgcaagctcttctgccaacataaattcattataaacaaga 3903  
|||||  
Db 481 TTGAACAGAGAGGAGATTGCAAGATCTTATGCCAACATAATTCATTAATAAACAGA 540  
|||||  
QY 3904 aaacaggtattgcacagttggtgaagta tggcttaaaagaccatagagggtgtgtgac 3963  
|||||  
Db 541 AAACAGGTATTGCACAGTTGGTGAAGTATGGCTTAAAGACCTAGAGGAGTGTGGAC 600  
|||||  
QY 3964 tgtgaagaactcggcatcaagttacaggtcttgatcaatttggttggcttcaag 4023  
|||||  
Db 601 TGTGAAGAAACTCGGCATCAAGTTACAGTCTTGATCAATTTGGCTTTGGTTTACAAG 560  
|||||  
QY 4024 tgcagcagcaaatggaatcattctccagttgtgtgctttcatcaaacgaaggaaggg 4083  
|||||  
Db 661 TGCAGCAGCAATGGAATCATCTCCAGTTTGTGGCTTTCATCAAAACGAAGCAAGGG 720  
|||||  
QY 4084 ctgtacctgaaatctctgacgtggtgaggtgaggtatgacgtgtgattccccagtttag 4143  
|||||  
Db 721 CTGTACCTGAAATCTCGCAGCTGGAGGAGATATGACCTGTGATTCCTCCAGTTTAGAG 780  
|||||  
QY 4144 ggcacagctctgggcccagttcccaactgcttgggtcagcagtagctatagacaaga 4203  
|||||  
Db 781 GGCCACAGCTCTGGGCCAAGTTCCCACTGCCATTTGGGTACACATAGCTATAGACAAGA 840  
|||||  
QY 4204 tatctgtctgtctcaaatgaggaatctgtttacaataaagcttctgtgacctctgg 4263  
|||||  
Db 841 TATCTGCTGTCTCTCAACATGGAGGAATCTGTTACAATAAGCTCTTGTGACCTCTGG 900  
|||||  
QY 4264 ttgtaagtgtgtcagatgtctatgtccagggccatcaactcaaccagaaactctga 4323  
|||||  
Db 901 TTGTAAAGTGTGGTCAAGATGCTATGTGCCAGGGCCATCAACCTTAACCCAGAAATCTCGA 960  
|||||  
QY 4324 cagcagcattcacagcagaatcatgtacagctgggtcacagtcaccaagaggaattacaag 4383  
|||||  
Db 961 CAGCAGGCATCACAGCAGAATCATGTACGACTGGTTCACAGTCCCAAGAGGAATTAACAG 1020  
|||||  
QY 4384 agtactgcagacatcatgaaatcactatgtggtccctgtctcggataaagaaggaagcc 4443  
|||||  
Db 1021 AGTACTGCAGACATCATGAATCACCCTATGTGGCCCTTGTCTCGGATTAAGAAGGAAGCC 1080  
|||||  
QY 4444 atgtcaaggttaagctcttcagagaagaagggcagacagagagcgtgtgtgagactg 4503  
|||||  
Db 1081 ATGTCAAGGTTAAGTCTTTCGAGAAGGAAGGCAAGCAGACAGAGCGTGTCTGGAGACTG 1140  
|||||  
QY 4504 aactgtgacactgtactgcagaaactgaggactaaagtcactgtatgaaggaatggca 4563  
|||||  
Db 1141 AACTTGTGACCATGTACTGCAGAACTAGAGACTAAAGTCACTGATGAAGGAATGGCA 1200  
|||||  
QY 4564 gagaagcttcgataattcttcagtgcaaaatctgaaaggggtcaatttctaagcttcag 4623  
|||||  
Db 1201 GAGAAGCTTCCGTAATCTTCAGTGCAGAAATCTGAAGGGGTCAATTTCTTAATGCTTCAG 1260  
|||||  
QY 4624 gttgtttgaaatccatgagcaacagtggttccattgtgagtggtgtagcccccggaga 4683  
|||||  
Db 1261 GTTTGTTTGAATTCATCGAGCAACAGTGGTTCCTCATGTGTGAGTGTGACCCCGGAGA 1320  
|||||  
QY 4684 agctgcagcagcactaggaggcgtatgaaactcaggtgataaactcagactcagacct 4743  
|||||  
Db 1321 AGCTGTACGACGACTAGGAGGCGCTATGAACTCAGGTCAAACTCGACTTCAGACCT 1380  
|||||  
QY 4744 ccttgcaacttaacatcagaaagcagtgaaattgaaattcgtgtggtatcaccca 4803  
|||||  
Db 1381 CCTTTGCCAACTTACATCAGAAAGCAGTGAAATTTGAAATTTCTGGCTGTGGATCTACCCA 1440  
|||||

QY 4804 aagaacaataattacactttttatcattacagtgaggctgatgaacagcgcatttaaca 4863  
|||||  
Db 1441 AAGAAACAATATTACAGTTTATCATTAGAGTGGATGCTGATGAACACAGCATTTAACA 1500  
|||||  
QY 4864 caactgtgaagcagctgctgtcacgcctgccaagcgaagatacctcaaatattagttctgtg 4923  
|||||  
Db 1501 CAACTGTGAACGACGCTGCTGTCAACGCTGCCAAAGCAAGATACCTCAAAATAGTCTGTG 1560  
|||||  
QY 4924 atgaattataacatcaagtagaataaaggtgtctgtgtctatttctgtacagctata 4983  
|||||  
Db 1561 ATGAATTTATTAACATCAAGTAGAAAAAGGTGTCTGTCTATTTCTGTACAGCTATA 1620  
|||||  
QY 4984 gagatgactactacagaatttatttaaccctaaagaactgtgttaacctcattcaaa 5043  
|||||  
Db 1621 GAGATGACTACTACAGATCTTATTTAAACCTTAAGAACTGTCTGTTAACTCATTTCAA 1680  
|||||  
QY 5044 cagacagaggttactactggaataatggaattgtgtacattcattcaataatttaataa 5103  
|||||  
Db 1681 CAGACAGAGGCTTATACCTGGAATATGGAATGTTGTACATTCATATAATTTAAATTA 1740  
|||||  
QY 5104 attctaaagaagctgggtgagtgctcacacttaacccagcacttttgggaagcc 5163  
|||||  
Db 1741 ATTCTAAGAAGAGGCTGGTGCAGTGGCTCACACCTTTAATCCAGCAGCTTTGGGAAGCC 1800  
|||||  
QY 5164 aaggcaggaagactgcttgaacccagggagtttgagaccagcctgagcaacaagaagac 5223  
|||||  
Db 1801 RAGGCAGGAAGACTGCTTGAACCAAGGAGTTTGAGACCAGCTTGAGCAACAAAGCAAGAC 1860  
|||||  
QY 5224 cccatctctataaaacttaaaaaattagttgggcattgtggcagctgcctgtagtccca 5283  
|||||  
Db 1861 CCCATCTCTATAAAAAACTTAAAAAAATTAGTTGGGCATGGTGACATGCCCTGTAGTCCCA 1920  
|||||  
QY 5284 gctactccagagctgagatgagatcatctgagcctcagaggttgaggtgagctgagct 5343  
|||||  
Db 1921 GCTACTCCAGAGCTGAGATGGATCATCTGACCTCAGAGGTTGAGGCTGCAGTGAGCT 1980  
|||||  
QY 5344 gtgactgagccactgcactccagctgagcaacaagcagcagcctgctcttaaaaaaaa 5403  
|||||  
Db 1981 GTGACTGCCCACTGCCTCCAGTCTGGGACAAACAGAGCAGACCCCTGTCTTTAAAAA 2040  
|||||  
QY 5404 aaagaaaaaaatttttttctaaagaagctgctcctcaaaagttgagctttgttagtttt 5463  
|||||  
Db 2041 AAAGAAAAAAATTTTCTTAAGAGCTGTCTTACAAAGTTGAGCTTTGTTAGTTT 2100  
|||||  
QY 5464 tcatgtgataatattataaatttatcttttgggataataaaatgctttcattacactg 5523  
|||||  
Db 2101 TCATGTGTAATATATATAAATTTATCTTTGGGATATAATAAATGCTTTCATATACCTG 2160  
|||||  
QY 5524 ca 5525  
||  
Db 2161 CA 2162  
||

## RESULT 10

HSM802022

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1917 bp mRNA PRI 18-FEB-2000  
Homo sapiens mRNA; cDNA DKFp434H149 (from clone DKFp434H149).  
AL137676  
GI:6807846  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1917)  
Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
Direct Submission  
Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152  
Martinsried, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.  
This clone (DRF2p43H149) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

## FEATURES

Location/Qualifiers  
1..1917  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DRF2p43H149"  
/clone\_lib="434 (synonym: htes3). Vector pSport1; host  
DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
/tissue\_type="testis"  
polyA\_signal  
1891..1896  
polyA\_site  
1907  
BASE COUNT 612 a 376 c 416 g 513 t  
ORIGIN

Query Match 30.5%; Score 1687.6; DB 93; Length 1917;  
Best Local Similarity 91.1%; Pred. No. 0;  
Matches 1903; Conservative 0; Mismatches 4; Indels 183; Gaps 3;

QY 3431 aagatatgtggcaagaataatattgaatttaaaacgatactgcatagaacgtgtgt 3490  
|||||  
Db 1 AAGATATGTGGCAAGAAATAATATTGAATTTAAACGATACGTCATAGAACGTGTGT 60  
|||||

QY 3491 caggccgcgaagttagatcatcttcccaagaacttcggagtgatgcattgtat 3550  
|||||  
Db 61 CAGGCGGCGCAAGTTAGATTCATTTCATCCCAAGAACTTCGAGGTGTGCATTGTAT 120  
|||||

QY 3551 tgtcacttaccacaaacagcttcttgcacactgctgaaattatctacactatctatga 3610  
|||||  
Db 121 TGTCACTTCTACCAACACAGCTTTCGCCACTGCTGAAATTAATCTACACTATCTATGA 180  
|||||

QY 3611 aatcatcaagagttccagcaacttcaggaaagaataacagttatttattgaacctac 3670  
|||||  
Db 181 AATCATCCAAAGAGTTTCAGCACCTTCAGGAAAGAAATTAACAGTATTATTGTAACCATAC 240  
|||||

QY 3671 catgttattgaagaacttcttaccactgtggatcccgagagataaaactcagtcagt 3730  
|||||  
Db 241 CATGTTATTGAAGCAATACTCTTACACTGTGGGATCCGAGAGATAAATCTCAGTCAAGT 300  
|||||

QY 3731 ctacattattctgtatgctgtgacagagaagctgacgagagagaagtggaaagctaa 3790  
|||||  
Db 301 CTACATTATTCTGTATGCTGTGACAGAGAAGCTGACGAGGAGAGAAAGTGGAAAGCTAA 360  
|||||

QY 3791 atttgttaactctgttcttcttaatagtctgtgctgactctacaaagtatttgaaca 3850  
|||||  
Db 361 ATTTGTGAATCTGCTTTGTCTCTAATACTGTCTGTGCGACTCTACAAGTTTATTGAACA 420  
|||||

QY 3851 gaaggagatttgcagatcttattgccaacaataaaattcattataaaacagaaaaacagg 3910  
|||||  
Db 421 GAAGGAGATTTCAGATCTTATGCCAACAAATAAATTCATTATTAACAGAAAAACAGG 480  
|||||

QY 3911 tattgcacagtttggatggctttaaagacacctagagaggttggactgttgaa 3970  
|||||  
Db 481 TATTGCAAGTTTGGTGAAGTATGGCTTTAAAGACCTAGAGAGAGTGTGGACTGTGAA 540  
|||||

QY 3971 gaaactcgcatcaagttacagcttctgattcaatttggcttggtttacaaggcgagca 4030  
|||||  
Db 541 GAAACTCTGCATCAAGTTACAGGCTCTGATCAATTTGGGCTTGTGTTTACAAGGTGCAGCA 600  
|||||

QY 4031 gcaaatgggaatcatcttccagtttggctttcacaacagcaaggcaagggtgtacc 4090  
|||||  
Db 601 GCAAAATGGAATCATCTCCAGTTTGTGGCTTTTCATCAAAAGCAAGGCAAGGCTGTACC 560  
|||||

QY 4091 tgaactctcgagctggagcgagatgatgacctgtgattccccagtttagaggggcaca 4150  
|||||  
Db 661 TGAATCTCTCGCAGCTGGAGGCAGATATACCTGCTGATTCCCCAGTTTAGAGGGGCCACA 720  
|||||

QY 4151 agctctggggcccaattcccaactgcccattgggggtcagcatagctatagacaagaatatctgc 4210  
|||||  
Db 721 AGCTCTGGGGCCAGTTCCCACTGCCATTGGGGTCCAGCATAGCTATAGACAAGATATCTGC 780  
|||||

QY 4211 tgcgtctctcaaatgaggaatctgttacaataaagctcttctgacccctctgtttaaag 4270  
|||||  
Db 781 TGCCTGTCTCAACATGGAGGAATCTGTACATAAGCTCTTTGTGACCTCTCTGTGTTGAAG 840  
|||||

QY 4271 tgttgggtcagatgtctatgtccaggggccatcaacacccaaacacacacacacacacag 4330  
|||||  
Db 841 TGTGGCCAGATGTCTATGTCCAGGGCCATCAACCTAAACCCAGAACTCTGGACACAGCAGG 900  
|||||

QY 4331 catcacagcagaatcatgtacacactggtccacagtcacagtcacacagagaattacaagagtag 4390  
|||||  
Db 901 CATCACAGCAGAATCATGTACGACTGGTCACAGTCCCAAGAGGAATTCACAAAGTACTG 960  
|||||

QY 4391 cagacatcatgaatacacctatgtgccccttctctcggaataaagaagaagacacatgtcaa 4450  
|||||  
Db 961 CAGACATCATGAATCAGCTATGTGGCCCTTGTCTCGGATAAAGAGGAGCCATGTCA- 1019  
|||||

QY 4451 ggttaagctcttcgagaagaagaaggcagacagagacgctgtgctggagactgaactgt 4510  
|||||  
Db 1020 ----- 1019  
|||||

QY 4511 ggaccatgtactgcagaataactgaggactaaagtcactgatgaaaggaatggcagagaagc 4570  
|||||  
Db 1020 -----AGAGAGC 1027  
|||||

QY 4571 ttcggataatcttgcagtgcaaaatctgaagggttcattttcttaattgcttcaggtttgtt 4630  
|||||  
Db 1028 TTCCGATAATCTTGCAGTGCMAAATCTGAAGGGTTCATTTCTAATGCTTTCAGGTTGTT 1087  
|||||

QY 4631 tgaataatccatggagcaaacagtggttcccaattgtgagtgctgtagcccccagagaagctgc 4690  
|||||  
Db 1088 TGAATTCATGGAGCAACAGTGGTTCCTCATTTGTGAGTGTGCTAGCCCGGAGAGCTGTC 1147  
|||||

QY 4691 agccagcacttagggcgctatgaaactcaggtacaaaactcgacttcagacccctctgc 4750  
|||||  
Db 1148 AGCCAGCACTAGGAGCGCTATGAACCTCAG----- 1178  
|||||

QY 4751 caacttacatcgaaaaagcagtgaaattctgctgtggtatctaccccaagaac 4810  
|||||  
Db 1179 -----GTGGATCTACCCAAAGAAC 1198  
|||||

QY 4811 aatattacagtttttatcattagatggagtgatcgtatgaacacagcatttaacacacactgt 4870  
|||||  
Db 1199 AATATTACAGTTTTTTATCATTTAGATGGGATCTGATGAACAGGCATTTAACACAACTGT 1258  
|||||

QY 4871 gaagcagctgctgtcacgcctgccaaagcaagataacctcaaatagctctgtatgaat 4930  
|||||  
Db 1259 GAAGCAGCTGTCTCAGCGCTGCCAAAGCAAGATACCTCAAAATAGTCTGTGATGAAT 1318  
|||||

QY 4931 ttataacatcaagtagaaaaaagggtctgtgtgtatattctgtacagctatagatga 4990  
|||||  
Db 1319 TTATAACATCAAAAGTAGAAAAAAGGTGTCTGTCTATTCTGTACAGCTATAGAGATGA 1378  
|||||

QY 4991 ctactacagaatcttatittaacctcaagaactgcttaacctcattcaaacacagacag 5050  
|||||  
Db 1379 CTACTACAGATCTTATTTTAAACCTTAAGAACTGTCTGTAACTCATTTCAACACAGACAG 1438  
|||||

QY 5051 aggcttacttggaaataatgggaattgttaccatcatcaataatttaaaattctaa 5110  
|||||  
Db 1439 AGGCTTACTTGAATAATATGGAATTTGTACATCATCAATTAATTTAAATTTAAATTTCTAA 1498  
|||||

QY 5111 gaaagcgtgggtgcagtggtccacacatttaaccacagactttgggaagccagagcag 5170  
|||||  
Db 1499 GAAGAGGCTGGGTGCGAGTGGCTCACACCTTTAATCCAGACCTTTTGGGAAGCCAGAGCAG 1558  
|||||

QY 5171 gaagactgcttgaacaccaggtgtttgagaccagcctgagcaacaaagcaagacccccatct 5230  
|||||  
Db 1559 GAAGACTGCTTGAACACAGGAGTTTGAGACCAGCTTGAGCAACAAAGCAAGACCCCATCT 1618  
|||||

QY 5231 ctataaaactaaaaaattagttgggcagatggtggcacatgctgtagtcocagctactc 5290  
|||||









All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L3086  
Center clone name: 10\_K15

----- Summary Statistics

Sequencing vector: M13; M7815; 98% of reads  
Sequencing vector: Plasmid; n/a; 8-0.1% of reads  
1.91489361702128Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731  
Consensus quality: 142640 bases at least Q40  
Consensus quality: 152432 bases at least Q30  
Consensus quality: 157096 bases at least Q20  
Insert size: 153000; agarose-fp  
Quality coverage: 160997; sum-of-contigs  
Quality cov. 3.4 in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 34 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1 1064: contig of 1064 bp in length
* 1085 1164: gap of 100 bp
* 1165 2366: contig of 1202 bp in length
* 2367 2466: gap of 100 bp
* 2467 3540: contig of 1074 bp in length
* 3541 3640: gap of 100 bp
* 3641 4861: contig of 1221 bp in length
* 4862 4961: gap of 100 bp
* 4962 5968: contig of 1007 bp in length
* 5969 6068: gap of 100 bp
* 6069 7242: contig of 1174 bp in length
* 7243 7342: gap of 100 bp
* 7343 8433: contig of 1111 bp in length
* 8434 8553: gap of 100 bp
* 8554 9820: contig of 1267 bp in length
* 9821 9920: gap of 100 bp
* 9921 10938: contig of 1018 bp in length
* 10939 11038: gap of 100 bp
* 11039 12256: contig of 1218 bp in length
* 12257 12356: gap of 100 bp
* 12357 14248: contig of 1892 bp in length
* 14249 14348: gap of 100 bp
* 14349 15680: contig of 1332 bp in length
* 15681 15780: gap of 100 bp
* 15781 17240: contig of 1460 bp in length
* 17241 17340: gap of 100 bp
* 17341 18693: contig of 1353 bp in length
* 18694 18793: gap of 100 bp
* 18794 20812: contig of 2019 bp in length
* 20813 20912: gap of 100 bp
* 20913 24906: contig of 3994 bp in length
* 24907 25006: gap of 100 bp
* 25007 28926: contig of 3920 bp in length
* 28927 29026: gap of 100 bp
* 29027 34159: contig of 5133 bp in length
* 34160 34259: gap of 100 bp
* 34260 37975: contig of 3716 bp in length
* 37976 38075: gap of 100 bp
* 38076 43955: contig of 5880 bp in length
* 43956 44055: gap of 100 bp
* 44056 48039: contig of 3984 bp in length

```

```

* 48040 48139: gap of 100 bp
* 48140 53767: contig of 5628 bp in length
* 53768 53967: gap of 100 bp
* 53968 59812: contig of 5945 bp in length
* 59813 59912: gap of 100 bp
* 59913 63066: contig of 3154 bp in length
* 63067 63166: gap of 100 bp
* 63167 70627: contig of 7461 bp in length
* 70628 70727: gap of 100 bp
* 70728 77701: contig of 6974 bp in length
* 77702 77801: gap of 100 bp
* 77802 85553: contig of 7752 bp in length
* 85554 85653: gap of 100 bp
* 85654 94602: contig of 8949 bp in length
* 94603 94702: gap of 100 bp
* 94703 104093: contig of 9391 bp in length
* 104094 104193: gap of 100 bp
* 104194 112395: contig of 8202 bp in length
* 112396 112495: gap of 100 bp
* 112496 123244: contig of 10749 bp in length
* 123245 123344: gap of 100 bp
* 123345 135427: contig of 12083 bp in length
* 135428 135527: gap of 100 bp
* 135528 147434: contig of 11907 bp in length
* 147435 147534: gap of 100 bp
* 147535 164297: contig of 16763 bp in length.

```

#### FEATURES

```

source
1..164297
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-10K15"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature
1..1064
/note="assembly_fragment"
1165..2366
/note="assembly_fragment"
2467..3540
/note="assembly_fragment"
3641..4861
/note="assembly_fragment"
4962..5968
/note="assembly_fragment"
6069..7242
/note="assembly_fragment"
7343..8453
/note="assembly_fragment"
8554..9820
/note="assembly_fragment"
9921..10938
/note="assembly_fragment"
11039..12256
/note="assembly_fragment"
12357..14248
/note="assembly_fragment"
14349..15680
/note="assembly_fragment"
15781..17240
/note="assembly_fragment"
17341..18693
/note="assembly_fragment"
18794..20812
/note="assembly_fragment"
20913..24906
/note="assembly_fragment"
25007..28926
/note="assembly_fragment"
29027..34159
/note="assembly_fragment"
34260..37975
/note="assembly_fragment"
clone_end:T7
vector_side:left"
38076..43955

```





Db 48029 AGCCTGAGCAACAAAGCAAGACCCCATCTCTATAAACTAAAAAATTANTTGGCATG 48088  
QY 5262 gtggcacatgctgtagtagtccagctactccagagctgagatgagatcatctgagcctcag 5321  
|||||  
Db 48089 GTGGCACATGCTGTAGTCCAGCTACTCCAGAGGCTGAGATGGATCATCTGAGCCTCAG 48148  
QY 5322 gaggttgaggtgcagtgagctgtgactgcgccactgcactccagctctgggacaacagag 5381  
|||||  
Db 48149 GAGGTTGAGGCTGCAGTGAGCTGTGACTGGCCCACTGCCACTCCAGCTCGGGACAACAGAG 48208  
QY 5382 caagaccctgtcttaaaaaaaagaaaaaaatTTTTTcttaagaagctgtccctac 5441  
|||||  
Db 48209 CAAGACCCTGTCTTAAAAAAGAAAAAATTTTTTCTAAGAAGCTGTCCTAC 48268  
QY 5442 aaagttgagctttagtttttcag 5468  
|||||  
Db 48269 AAAAGTTGAGCTTTGTAGTTTTTCATG 48295

Search completed: October 21, 2001, 18:18:05  
Job time: 26587 sec

